

B2

(19)



Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11)

EP 0 619 321 B1

(12)

## EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention  
of the grant of the patent:  
07.01.1999 Bulletin 1999/01

(51) Int Cl.<sup>6</sup>: C07K 1/04, C07K 17/06,  
C07K 17/14, B01J 19/00,  
G03F 7/38, C07H 21/00

(21) Application number: 94200059.7

(22) Date of filing: 07.06.1990

## (54) Method and apparatus for investigating polynucleotide or amino acid sequences

Verfahren und Vorrichtung zur Untersuchung von Polynukleotide oder Aminosäure Sequenzen

Méthode et appareil pour l'investigation des sequences polynucleotides ou des acides aminés

(84) Designated Contracting States:  
AT BE CH DE DK ES FR GB IT LI LU NL SE

(56) References cited:  
EP-A- 0 328 256

(30) Priority: 07.06.1989 US 362901  
07.03.1990 US 492462

(43) Date of publication of application:  
12.10.1994 Bulletin 1994/41

(60) Divisional application: 98203518.0

(62) Document number(s) of the earlier application(s) in  
accordance with Art. 76 EPC:  
90909187.8 / 0 476 014

(73) Proprietor: AFFYMAX TECHNOLOGIES N.V.  
Willemstad, Curaçao (AN)

(72) Inventors:  
• Pirrung, Michael C.  
Durham NC 27707 (US)  
• Read, J. Leighton  
Palo Alto CA 94301 (US)  
• Fodor, Stephen P.A.  
Palo Alto CA 94303 (US)  
• Stryer, Lubert  
Stanford CA 94305 (US)

(74) Representative: Bizley, Richard Edward et al  
Hepworth, Lawrence, Bryer & Bizley  
Merlin House  
Falconry Court  
Baker's Lane  
Epping Essex CM16 5DQ (GB)

- CHEMICAL ABSTRACTS, vol. 110, no. 9, 27 February 1989, Columbus, Ohio, US; abstract no. 76031w, V K HARIDASAN ET AL. 'Peptide synthesis using photolabilely cleavable 2-nitrobenzyloxycarbonyl protecting group' page 707 ; & PROC. INDIAN. NATL. SCI. ACAD., PART A vol. 53, no. 6 , 1987 pages 717 - 728
- CHEMICAL ABSTRACTS, vol. 100, no. 17, 23 April 1984, Columbus, Ohio, US; abstract no. 139591v, W STUEBER ET AL. 'synthesis and photolytic cleavage of bovine insulin B22-30 on a nitrobenzoylglycyl-poly(ethylene glycol) support' page 700 ; & INTERNATIONAL JOURNAL OF PEPTIDE AND PROTEIN RESEARCH vol. 22, no. 3 , 1983 pages 277 - 283
- CHEMICAL ABSTRACTS, vol. 93, no. 22, 1 December 1980, Columbus, Ohio, US; abstract no. 213252r, M GAZARD ET AL. 'lithographic technique using radiation-induced grafting of acrylic acid into poly (methyl methacrylate) films' page 585 ; & POLYM. ENG. SCI vol. 20, no. 16 , 1980 pages 1069 - 1072
- SYNTHESIS, no. 1, 1980, STUTTGART DE pages 1 - 26 V N RAJASEKHARAN PILLAI 'photoremovable protecting groups in organic synthesis'
- SCIENCE, vol. 251 , 15 February 1991 , LANCASTER, PA US pages 767 - 773 S P A FODOR ET AL. 'light-directed, spatially addressable parallel chemical synthesis'

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 0 619 321 B1

## Description

## BACKGROUND OF THE INVENTION

5 The present invention relates to a method and apparatus for investigating polynucleotide or amino acid sequences. The apparatus comprises diverse chemical sequences at known locations on a single substrate surface.

The relationship between structure and activity of molecules is a fundamental issue in the study of biological systems. Structure-activity relationships are important in understanding, for example, the function of enzymes, the ways in which cells communicate with each other, as well as cellular control and feedback systems.

10 Certain macromolecules are known to interact and bind to other molecules having a very specific three-dimensional spatial and electronic distribution. Any large molecule having such specificity can be considered a receptor, whether it is an enzyme catalyzing hydrolysis of a metabolic intermediate, a cell-surface protein mediating membrane transport of ions, a glycoprotein serving to identify a particular cell to its neighbors, an IgG-class antibody circulating in the plasma, an oligonucleotide sequence of DNA in the nucleus, or the like. The various molecules which receptors selectively bind are known as ligands.

15 Many assays are available for measuring the binding affinity of known receptors and ligands, but the information which can be gained from such experiments is often limited by the number and type of ligands which are available. Novel ligands are sometimes discovered by chance or by application of new techniques for the elucidation of molecular structure, including x-ray crystallographic analysis and recombinant genetic techniques for proteins.

20 Small peptides are an exemplary system for exploring the relationship between structure and function in biology. A peptide is a sequence of amino acids. When the twenty naturally occurring amino acids are condensed into polymeric molecules they form a wide variety of three-dimensional configurations, each resulting from a particular amino acid sequence and solvent condition. The number of possible pentapeptides of the 20 naturally occurring amino acids, for example, is  $20^5$  or 3.2 million different peptides. The likelihood that molecules of this size might be useful in receptor-binding studies is supported by epitope analysis studies showing that some antibodies recognize sequences as short as a few amino acids with high specificity. Furthermore, the average molecular weight of amino acids puts small peptides in the size range of many currently useful pharmaceutical products.

25 Pharmaceutical drug discovery is one type of research which relies on such a study of structure-activity relationships. In most cases, contemporary pharmaceutical research can be described as the process of discovering novel ligands with desirable patterns of specificity for biologically important receptors. Another example is research to discover new compounds for use in agriculture, such as pesticides and herbicides.

30 Sometimes, the solution to a rational process of designing ligands is difficult or unyielding. Prior methods of preparing large numbers of different polymers have been painstakingly slow when used at a scale sufficient to permit effective rational or random screening. For example, the "Merrifield" method (*J. Am. Chem. Soc.* (1963) **85**:2149-2154, which is incorporated herein by reference for all purposes) has been used to synthesize peptides on a solid support. In the Merrifield method, an amino acid is covalently bonded to a support made of an insoluble polymer. Another amino acid with an alpha protected group is reacted with the covalently bonded amino acid to form a dipeptide. After washing, the protective group is removed and a third amino acid with an alpha protective group is added to the dipeptide. This process is continued until a peptide of a desired length and sequence is obtained. Using the Merrifield method, it is not economically practical to synthesize more than a handful of peptide sequences in a day.

40 To synthesize larger numbers of polymer sequences, it has also been proposed to use a series of reaction vessels for polymer synthesis. For example, a tubular reactor system may be used to synthesize a linear polymer on a solid phase support by automated sequential addition of reagents. This method still does not enable the synthesis of a sufficiently large number of polymer sequences for effective economical screening.

45 Methods of preparing a plurality of polymer sequences are also known in which a foraminous container encloses a known quantity of reactive particles, the particles being larger in size than foramina of the container. The containers may be selectively reacted with desired materials to synthesize desired sequences of product molecules. As with other methods known in the art, this method cannot practically be used to synthesize a sufficient variety of polypeptides for effective screening.

50 Other techniques have also been described. These methods include the synthesis of peptides on 96 plastic pins which fit the format of standard microtiter plates. Unfortunately, while these techniques have been somewhat useful, substantial problems remain. For example, these methods continue to be limited in the diversity of sequences which can be economically synthesized and screened.

55 From the above, it is seen that an improved method and apparatus for synthesizing a variety of chemical sequences at known locations is desired. Such a method and apparatus is described and claimed in European Patent No 0 476 014, from which this case is divided.

## SUMMARY OF THE INVENTION

In one aspect the present invention provides a method of investigating by receptor/ligand binding a polynucleotide or amino acid sequence by the use of a substrate with a surface, said surface comprising at least  $10^3$  predefined regions, said predefined regions containing different nucleotide or amino acid sequences thereon, said predefined regions each occupying an area of less than about  $2.5 \times 10^{-3} \text{ cm}^2$ , which method comprises labelling said sequence being investigated and identifying which of said different sequences binds with said sequence being analysed.

In a further aspect the invention provides apparatus for investigating by receptor/ligand binding a polynucleotide or amino acid sequence comprising a substrate with a surface, said surface comprising at least  $10^3$  predefined regions, said predefined regions containing different polynucleotide or amino acid sequences thereon, said predefined regions each occupying an area of less than about  $2.5 \times 10^{-3} \text{ cm}^2$ .

The invention further includes the use of an apparatus as defined above in investigating by receptor/ligand binding a polynucleotide or amino acid sequence.

In describing the present invention, it is first convenient to discuss the improved method and apparatus for the preparation of a variety of polymers of EP-B-0 476 014.

Preferably, linker molecules are provided on a substrate. A terminal end of the linker molecules is provided with a reactive functional group protected with a photoremovable protective group. Using lithographic methods, the photoremovable protective group is exposed to light and removed from the linker molecules in first selected regions. The substrate is then washed or otherwise contacted with a first monomer that reacts with exposed functional groups on the linker molecules. The monomer may be an amino acid containing a photoremovable protective group at its amino or carboxy terminus and the linker molecule terminates in an amino or carboxy acid group bearing a photoremovable protective group.

A second set of selected regions is, thereafter, exposed to light and the photoremovable protective group on the linker molecule/protected amino acid is removed at the second set of regions. The substrate is then contacted with a second monomer containing a photoremovable protective group for reaction with exposed functional groups. This process is repeated to selectively apply monomers until polymers of a desired length and desired chemical sequence are obtained. Photolabile groups are then optionally removed and the sequence is, thereafter, optionally capped. Side chain protective groups, if present, are also removed.

By using the lithographic techniques disclosed herein, it is possible to direct light to relatively small and precisely known locations on the substrate. It is, therefore, possible to synthesize polymers of a known chemical sequence at known locations on the substrate.

The resulting substrate will have a variety of uses including, for example, screening large numbers of polymers for biological activity. To screen for biological activity, the substrate is exposed to one or more receptors such as antibody whole cells, receptors on vesicles, lipids, or any one of a variety of other receptors. The receptors are preferably labeled with, for example, a fluorescent marker, radioactive marker, or a labeled antibody reactive with the receptor. The location of the marker on the substrate is detected with, for example, photon detection or autoradiographic techniques. Through knowledge of the sequence of the material at the location where binding is detected, it is possible to quickly determine which sequence binds with the receptor and, therefore, the technique can be used to screen large numbers of peptides. Other possible applications of the invention include diagnostics in which various antibodies for particular receptors would be placed on a substrate and, for example, blood sera would be screened for immune deficiencies.

An improved reactor system for synthesizing polymers is also disclosed in EP-B-0 476 014. The reactor system includes a substrate mount which engages a substrate around a periphery thereof. The substrate mount provides for a reactor space between the substrate and the mount through or into which reaction fluids are pumped or flowed. A mask is placed on or focused on the substrate and illuminated so as to deprotect selected regions of the substrate in the reactor space. A monomer is pumped through the reactor space or otherwise contacted with the substrate and reacts with the deprotected regions. By selectively deprotecting regions on the substrate and flowing predetermined monomers through the reactor space, desired polymers at known locations may be synthesized.

Improved detection apparatus and methods are also disclosed. The detection method and apparatus utilize a substrate having a large variety of polymer sequences at known locations on a surface thereof. The substrate is exposed to a fluorescently labeled receptor which binds to one or more of the polymer sequences. The substrate is placed in a microscope detection apparatus for identification of locations where binding takes place. The microscope detection apparatus includes a monochromatic or polychromatic light source for directing light at the substrate, means for detecting fluoresced light from the substrate, and means for determining a location of the fluoresced light. The means for detecting light fluoresced on the substrate may in some embodiments include a photon counter. The means for determining a location of the fluoresced light may include an x/y translation table for the substrate. Translation of the slide and data collection are recorded and managed by an appropriately programmed digital computer.

A further understanding of the nature and advantages of the invention may be realized by reference to the remaining portions of the specification, and the attached drawings.

## BRIEF DESCRIPTION OF THE FIGURES

- Fig. 1 illustrates masking and irradiation of a substrate at a first location. The substrate is shown in cross-section;  
 Fig. 2 illustrates the substrate after application of a monomer "A";  
 Fig. 3 illustrates irradiation of the substrate at a second location;  
 Fig. 4 illustrates the substrate after application of monomer "B";  
 Fig. 5 illustrates irradiation of the "A" monomer;  
 Fig. 6 illustrates the substrate after a second application of "B";  
 Fig. 7 illustrates a completed substrate;  
 Figs. 8A and 8B illustrate alternative versions of a reactor system for forming a plurality of polymers on a substrate;  
 Fig. 9 illustrates a detection apparatus for locating fluorescent markers on the substrate;  
 Figs. 10A-10M illustrate the method as it is applied to the production of the trimers of monomers "A" and "B";  
 Figs. 11A, 11B, and 11C are fluorescence traces for standard fluorescent beads;  
 Figs. 12A and 12B are fluorescence curves for NVOC slides not exposed and exposed to light respectively;  
 Figs. 13A and 13B illustrate formation of a slide with a checkerboard pattern of YGGFL and GGFL exposed to labeled Herz antibody; and  
 Figs. 14A and 14B illustrate the mapping of sixteen sequences synthesized on two different glass slides.

## DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS CONTENTS

- I. Glossary  
 II. General  
 III. Polymer Synthesis  
 IV. Details of One Version of a Reactor System  
 V. Details of One Version of a Fluorescent Detection Device  
 VI. Determination of Relative Binding Strength of Receptors  
 VII. Examples  
 A. Slide Preparation  
 B. Synthesis of Eight Trimers of "A" and "B"  
 C. Synthesis of a Dimer of an Aminopropyl Group and a Fluorescent Group  
 D. Demonstration of Signal Capability  
 E. Determination of the Number of Molecules Per Unit Area  
 F. Removal of NVOC and Attachment of a Fluorescent Marker  
 G. Use of a Mask in Removal of NVOC  
 H. Attachment of YGGFL and Subsequent Exposure to Herz Antibody and Goat Antimouse  
 I. Monomer-by-Monomer Formation of YGGFL and Subsequent Exposure to Labeled Antibody  
 J. Monomer-by-Monomer Synthesis of YGGFL and PGGFL  
 K. Monomer-by-Monomer Synthesis of YGGFL and YPGGFL  
 L. Synthesis of an Array of Sixteen Different Amino Acid Sequences and Estimation of Relative Binding Affinity to Herz Antibody  
 VIII. Illustrative Alternative Approach  
 IX. Conclusion

## I. Glossary

The following terms are intended to have the following general meanings as they are used herein:

1. Complementary: Refers to the topological compatibility or matching together of interacting surfaces of a ligand molecule and its receptor. Thus, the receptor and its ligand can be described as complementary, and furthermore, the contact surface characteristics are complementary to each other.  
 2. Epitope: The portion of an antigen molecule which is delineated by the area of interaction with the subclass of receptors known as antibodies.  
 3. Ligand: A ligand is a molecule that is recognized by a particular receptor. Examples of ligands that can be investigated by this invention include, but are not restricted to, agonists and antagonists for cell membrane recep-

tors, toxins and venoms, viral epitopes, hormones (e.g., opiates, steroids, etc.), hormone receptors, peptides, enzymes, enzyme substrates, cofactors, drugs, lectins, sugars, oligonucleotides, nucleic acids, oligosaccharides, proteins, and monoclonal antibodies.

5       4. Monomer: A member of the set of small molecules which can be joined together to form a polymer. The set of monomers includes but is not restricted to, for example, the set of common L-amino acids, the set of D-amino acids, the set of synthetic amino acids, and the set of nucleotides and the set of pentoses and hexoses. As used herein, monomers refers to any member of a basis set for synthesis of a polymer. For example, dimers of L-amino acids form a basis set of 400 monomers for synthesis of polypeptides. Different basis sets of monomers may be used at successive steps in the synthesis of a polymer.

15       5. Peptide: A polymer in which the monomers are alpha amino acids and which are joined together through amide bonds and alternatively referred to as a polypeptide. In the context of this specification it should be appreciated that the amino acids may be the L-optical isomer or the D-optical isomer. Peptides are more than two amino acid monomers long, and often more than 20 amino acid monomers long. Standard abbreviations for amino acids are used (e.g., P for proline). These abbreviations are included in Stryer, Biochemistry, Third Ed., 1988, which is incorporated herein by reference for all purposes.

20       6. Radiation: Energy which may be selectively applied including energy having a wavelength of between  $10^{-14}$  and  $10^4$  meters including, for example, electron beam radiation, gamma radiation, x-ray radiation, ultraviolet radiation, visible light, infrared radiation, microwave radiation and radio waves. "Irradiation" refers to the application of radiation to a surface. microwave radiation, and radio waves. "Irradiation" refers to the application of radiation to a surface.

25       7. Receptor: A molecule that has an affinity for a given ligand. Receptors may be naturally-occurring or manmade molecules. Also, they can be employed in their unaltered state or as aggregates with other species. Receptors may be attached, covalently or noncovalently, to a binding member, either directly or via a specific binding substance. Examples of receptors which can be employed by this invention include, but are not restricted to, antibodies, cell membrane receptors, monoclonal antibodies and antisera reactive with specific antigenic determinants (such as on viruses, cells or other materials), drugs, polynucleotides, nucleic acids, peptides, cofactors, lectins, sugars, polysaccharides, cells, cellular membranes, and organelles. Receptors are sometimes referred to in the art as anti-ligands. As the term receptors is used herein, no difference in meaning is intended. A "Ligand Receptor Pair" is formed when two macromolecules have combined through molecular recognition to form a complex.

Other examples of receptors which can be investigated by this invention include but are not restricted to:

35       a) Microorganism receptors: Determination of ligands which bind to receptors, such as specific transport proteins or enzymes essential to survival of microorganisms, is useful in a new class of antibiotics. Of particular value would be antibiotics against opportunistic fungi, protozoa, and those bacteria resistant to the antibiotics in current use.

40       b) Enzymes: For instance, the binding site of enzymes such as the enzymes responsible for cleaving neurotransmitters; determination of ligands which bind to certain receptors to modulate the action of the enzymes which cleave the different neurotransmitters is useful in the development of drugs which can be used in the treatment of disorders of neurotransmission.

45       c) Antibodies: For instance, the invention may be useful in investigating the ligand-binding site on the antibody molecule which combines with the epitope of an antigen of interest; determining a sequence that mimics an antigenic epitope may lead to the development of vaccines of which the immunogen is based on one or more of such sequences or lead to the development of related diagnostic agents or compounds useful in therapeutic treatments such as for auto-immune diseases (e.g., by blocking the binding of the "self" antibodies).

d) Nucleic Acids: Sequences of nucleic acids may be synthesized to establish DNA or RNA binding sequences.

50       e) Catalytic Polypeptides: Polymers, preferably polypeptides, which are capable of promoting a chemical reaction involving the conversion of one or more reactants to one or more products. Such polypeptides generally include a binding site specific for at least one reactant or reaction intermediate and an active functionality proximate to the binding site, which functionality is capable of chemically modifying the bound reactant. Catalytic polypeptides are described in, for example, U.S. application Serial No. 404,920, which is incorporated herein by reference for all purposes.

55       f) Hormone receptors: For instance, the receptors for insulin and growth hormone. Determination of the ligands which bind with high affinity to a receptor is useful in the development of, for example, an oral replacement of the daily injections which diabetics must take to relieve the symptoms of diabetes, and in the other case, a

replacement for the scarce human growth hormone which can only be obtained from cadavers or by recombinant DNA technology. Other examples are the vasoconstrictive hormone receptors; determination of those ligands which bind to a receptor may lead to the development of drugs to control blood pressure.

g) Opiate receptors: Determination of ligands which bind to the opiate receptors in the brain is useful in the development of less-addictive replacements for morphine and related drugs.

8. Substrate: A material having a rigid or semi-rigid surface. In many embodiments, at least one surface of the substrate will be substantially flat, although in some embodiments it may be desirable to physically separate synthesis regions for different polymers with, for example, wells, raised regions, etched trenches, or the like. According to other embodiments, small beads may be provided on the surface which may be released upon completion of the synthesis.

9. Protective Group: A material which is bound to a monomer unit and which may be spatially removed upon selective exposure to an activator such as electromagnetic radiation. Examples of protective groups with utility herein include Nitroveratryloxy carbonyl, Nitrobenzyloxy carbonyl, Dimethyl dimethoxybenzyloxy carbonyl, 5-Bromo-7-nitroindolyl, *o*-Hydroxy- $\alpha$ -methyl cinnamoyl, and 2-Oxymethylene anthraquinone. Other examples of activators include ion beams, electric fields, magnetic fields, electron beams, x-ray, and the like.

10. Predefined Region: A predefined region is a localized area on a surface which is, was, or is intended to be activated for formation of a polymer. The predefined region may have any convenient shape, e.g., circular, rectangular, elliptical, wedge-shaped, etc. For the sake of brevity herein, "predefined regions" are sometimes referred to simply as "regions."

11. Substantially Pure: A polymer is considered to be "substantially pure" within a predefined region of a substrate when it exhibits characteristics that distinguish it from other predefined regions. Typically, purity will be measured in terms of biological activity or function as a result of uniform sequence. Such characteristics will typically be measured by way of binding with a selected ligand or receptor.

## II. General

The present invention is described by reference to substrates having a plurality of polymer sequences in predefined regions. The invention is described herein primarily with regard to molecules containing sequences of amino acids, but can readily be applied with both linear and cyclic polymers of nucleic acids, i.e. polynucleotides and with regard to peptides having either  $\alpha$ -,  $\beta$ - or  $\omega$ -amino acids.

The prepared substrate may, for example, be used in screening polymers as ligands for binding with a receptor, although it will be apparent that the invention could be used for the synthesis of a receptor for binding with a ligand. The substrate disclosed herein will have a wide variety of other uses. Merely by way of example, the invention herein can be used in determining peptide and nucleic acid sequences which bind to proteins, finding sequence-specific binding drugs, identifying epitopes recognized by antibodies, and evaluation of a variety of drugs for clinical and diagnostic applications, as well as combinations of the above.

The invention preferably envisages the use of a substrate "S" with a surface. Linker molecules "L" are optionally provided on a surface of the substrate. The purpose of the linker molecules, in some embodiments, is to facilitate receptor recognition of the synthesized polymers.

Optionally, the linker molecules may be chemically protected for storage purposes. A chemical storage protective group such as t-BOC (t-butoxycarbonyl) may be used in some embodiments. Such chemical protective groups would be chemically removed upon exposure to, for example, acidic solution and would serve to protect the surface during storage and be removed prior to polymer preparation.

On the substrate or a distal end of the linker molecules, a functional group with a protective group  $P_0$  is provided. The protective group  $P_0$  may be removed upon exposure to radiation, electric fields, electric currents, or other activators to expose the functional group.

Preferably, the radiation is ultraviolet (UV), infrared (IR), or visible light. As more fully described below, the protective group may alternatively be an electrochemically-sensitive group which may be removed in the presence of an electric field. Alternatively, ion beams, electron beams, or the like may be used for deprotection.

In preferred versions the exposed area to which each distinct polymer sequence is attached is less than about 10,000  $\mu\text{m}^2$  and may, in some versions encompass the binding site for as few as a single molecule. Within these regions, each polymer is preferably synthesized in a substantially pure form.

Concurrently or after exposure of a known region of the substrate to light, the surface is contacted with a first monomer unit  $M_1$  which reacts with the functional group which has been exposed by the deprotection step. The first

# EP 0 619 321 B1

monomer includes a protective group  $P_1$ .  $P_1$  may or may not be the same as  $P_0$ .

Accordingly, after a first cycle, known first regions of the surface may comprise the sequence:

5  $S-L-M_1-P_1$

while remaining regions of the surface comprise the sequence:

10  $S-L-P_0$

Thereafter, second regions of the surface (which may include the first region) are exposed to light and contacted with a second monomer  $M_2$  (which may or may not be the same as  $M_1$ ) having a protective group  $P_2$ .  $P_2$  may or may not be the same as  $P_0$  and  $P_1$ . After this second cycle, different regions of the substrate may comprise one or more of the following sequences:

15  $S-L-M_1-M_2-P_2$

20  $S-L-M_2-P_2$

25  $S-L-M_1-P_1$

and/or

30  $S-L-P_0$

The above process is repeated until the substrate includes desired polymers of desired lengths. By controlling the locations of the substrate exposed to light and the reagents exposed to the substrate following exposure, the location of each sequence will be known.

Thereafter, the protective groups are removed from some or all of the substrate and the sequences are, optionally, capped with a capping unit C. The process results in a substrate having a surface with a plurality of polymers of the following general formula:

35  $S-[L]-(M_1)-(M_2)-(M_3)...(M_x)-[C]$

40 where square brackets indicate optional groups, and  $M_1...M_x$  indicates any sequence of monomers. The number of monomers could cover a wide variety of values, but in a preferred embodiment they will range from 2 to 100.

In some embodiments a plurality of locations on the substrate polymers are to contain a common monomer subsequence. For example, it may be desired to synthesize a sequence  $S-M_1-M_2-M_3$  at first locations and a sequence  $S-M_4-M_2-M_3$  at second locations. The process would commence with irradiation of the first locations followed by contacting with  $M_1-P_1$ , resulting in the sequence  $S-M_1-P_1$  at the first location. The second locations would then be irradiated and contacted with  $M_4-P_1$ , resulting in the sequence  $S-M_4-P_1$  at the second locations. Thereafter both the first and second locations would be irradiated and contacted with the dimer  $M_2-M_3$ , resulting in the sequence  $S-M_1-M_2-M_3$  at the first locations and  $S-M_4-M_2-M_3$  at the second locations. Of course, common subsequences of any length could be utilized including those in a range of 2 or more monomers, 2 to 100 monomers, 2 to 20 monomers, and a most preferred range of 2 to 3 monomers.

According to other approaches, a set of masks is used for the first monomer layer and, thereafter, varied light wavelengths are used for selective deprotection. For example, in the process discussed above, first regions are first exposed through a mask and reacted with a first monomer having a first protective group  $P_1$ , which is removable upon exposure to a first wavelength of light (e.g., IR). Second regions are masked and reacted with a second monomer having a second protective group  $P_2$ , which is removable upon exposure to a second wavelength of light (e.g., UV). Thereafter, masks become unnecessary in the synthesis because the entire substrate may be exposed alternatively to the first and second wavelengths of light in the deprotection cycle.

The polymers prepared on a substrate according to the above methods will have a variety of uses including, for example, screening for biological activity. In such screening activities, the substrate containing the sequences is exposed to an unlabeled or labeled receptor such as an antibody, receptor on a cell, phospholipid vesicle, or any one of a variety of other receptors. In one preferred embodiment the polymers are exposed to a first, unlabeled receptor of interest and, thereafter, exposed to a labeled receptor-specific recognition element, which is, for example, an antibody. This process will provide signal amplification in the detection stage.

The receptor molecules may bind with one or more polymers on the substrate. The presence of the labeled receptor and, therefore, the presence of a sequence which binds with the receptor is detected in a preferred embodiment through the use of autoradiography, detection of fluorescence with a charge-coupled device, fluorescence microscopy, or the like. The sequence of the polymer at the locations where the receptor binding is detected may be used to determine all or part of a sequence which is complementary to the receptor.

Use of the invention herein is illustrated primarily with reference to screening for biological activity. The invention will, however, find many other uses. For example, the invention may be used in information storage (e.g., on optical disks), production of molecular electronic devices, production of stationary phases in separation sciences, production of dyes and brightening agents, photography, and in immobilization of cells, proteins, lectins, nucleic acids, polysaccharides and the like in patterns on a surface via molecular recognition of specific polymer sequences. By synthesizing the same compound in adjacent, progressively differing concentrations, a gradient will be established to control chemotaxis or to develop diagnostic dipsticks which, for example, titrate an antibody against an increasing amount of antigen. By synthesizing several catalyst molecules in close proximity, more efficient multistep conversions may be achieved by "coordinate immobilization." Coordinate immobilization also may be used for electron transfer systems, as well as to provide both structural integrity and other desirable properties to materials such as lubrication, wetting, etc.

According to alternative embodiments, molecular biodistribution or pharmacokinetic properties may be examined. For example, to assess resistance to intestinal or serum proteases, polymers may be capped with a fluorescent tag and exposed to biological fluids of interest.

### III. Polymer Synthesis

Fig. 1 illustrates one embodiment of the invention disclosed herein in which a substrate 2 is shown in cross-section. Essentially, any conceivable substrate may be employed in the invention. The substrate may be biological, nonbiological, organic, inorganic, or a combination of any of these, existing as particles, strands, precipitates, gels, sheets, tubing, spheres, containers, capillaries, pads, slices, films, plates, slides, etc. The substrate may have any convenient shape, such as a disc, square, sphere, circle, etc. The substrate is preferably flat but may take on a variety of alternative surface configurations. For example, the substrate may contain raised or depressed regions on which the synthesis takes place. The substrate and its surface preferably form a rigid support on which to carry out the reactions described herein. The substrate and its surface is also chosen to provide appropriate light-absorbing characteristics. For instance, the substrate may be a polymerized Langmuir Blodgett film, functionalized glass, Si, Ge, GaAs, GaP, SiO<sub>2</sub>, SiN<sub>4</sub>, modified silicon, or any one of a wide variety of gels or polymers such as (poly)tetrafluoroethylene, (poly)vinylidenedifluoride, polystyrene, polycarbonate, or combinations thereof. Other substrate materials will be readily apparent to those of skill in the art upon review of this disclosure. In a preferred embodiment the substrate is flat glass or single-crystal silicon with surface relief features of less than 10 Å.

According to some embodiments, the surface of the substrate is etched using well known techniques to provide for desired surface features. For example, by way of the formation of trenches, v-grooves, mesa structures, or the like, the synthesis regions may be more closely placed within the focus point of impinging light, be provided with reflective "mirror" structures for maximization of light collection from fluorescent sources, or the like.

Surfaces on the solid substrate will usually, though not always, be composed of the same material as the substrate. Thus, the surface may be composed of any of a wide variety of materials, for example, polymers, plastics, resins, polysaccharides, silica or silica-based materials, carbon, metals, inorganic glasses, membranes, or any of the above-listed substrate materials. In some embodiments the surface may provide for the use of caged binding members which are attached firmly to the surface of the substrate in accord with the teaching of copending application Serial No. 404,920, previously incorporated herein by reference. Preferably, the surface will contain reactive groups, which could be carboxyl, amino, hydroxyl, or the like. Most preferably, the surface will be optically transparent and will have surface Si-OH functionalities, such as are found on silica surfaces.

The surface 4 of the substrate is preferably provided with a layer of linker molecules 6, although it will be understood that the linker molecules are not required elements of the invention. The linker molecules are preferably of sufficient length to permit polymers in a completed substrate to interact freely with molecules exposed to the substrate. The linker molecules should be 6-50 atoms long to provide sufficient exposure. The linker molecules may be, for example, aryl acetylene, ethylene glycol oligomers containing 2-10 monomer units, diamines, diacids, amino acids, or combinations thereof. Other linker molecules may be used in light of this disclosure.



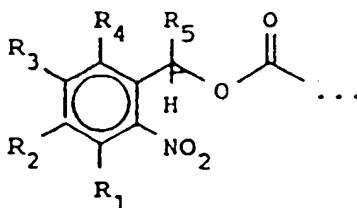
According to alternative embodiments, the linker molecules are selected based upon their hydrophilic/hydrophobic properties to improve presentation of synthesized polymers to certain receptors. For example, in the case of a hydrophilic receptor, hydrophilic linker molecules will be preferred so as to permit the receptor to more closely approach the synthesized polymer.

According to another alternative embodiment, linker molecules are also provided with a photocleavable group at an intermediate position. The photocleavable group is preferably cleavable at a wavelength different from the protective group. This enables removal of the various polymers following completion of the synthesis by way of exposure to the different wavelengths of light.

The linker molecules can be attached to the substrate via carbon-carbon bonds using, for example, (poly)trifluoroethoxyethylene surfaces, or preferably, by siloxane bonds (using, for example, glass or silicon oxide surfaces). Siloxane bonds with the surface of the substrate may be formed in one embodiment via reactions of linker molecules bearing trichlorosilyl groups. The linker molecules may optionally be attached in an ordered array, i.e., as parts of the head groups in a polymerized Langmuir Blodgett film. In alternative embodiments, the linker molecules are adsorbed to the surface of the substrate.

The linker molecules and monomers used herein are provided with a functional group to which is bound a protective group. Preferably, the protective group is on the distal or terminal end of the linker molecule opposite the substrate. The protective group may be either a negative protective group (i.e., the protective group renders the linker molecules less reactive with a monomer upon exposure) or a positive protective group (i.e., the protective group renders the linker molecules more reactive with a monomer upon exposure). In the case of negative protective groups an additional step of reactivation will be required. Sometimes, this will be done by heating.

The protective group on the linker molecules may be selected from a wide variety of positive light-reactive groups preferably including nitro aromatic compounds such as o-nitrobenzyl derivatives or benzylsulfonyl. Preferably, 6-nitroveratryloxy-carbonyl (NVOC), 2-nitrobenzyloxycarbonyl (NBOC) or  $\alpha,\alpha$ -dimethyl-dimethoxybenzyloxycarbonyl (DDZ) is used. A nitro aromatic compound containing a benzylic hydrogen ortho to the nitro group can be used, i.e., a chemical of the form:



where  $R_1$  is alkoxy, alkyl, halo, aryl, alkenyl, or hydrogen;  $R_2$  is alkoxy, alkyl, halo, aryl, nitro, or hydrogen;  $R_3$  is alkoxy, alkyl, halo, nitro, aryl, or hydrogen;  $R_4$  is alkoxy, alkyl, hydrogen, aryl, halo, or nitro; and  $R_5$  is alkyl, alkenyl, cyano, alkoxy, hydrogen, halo, aryl, or alkenyl. Other materials which may be used include o-hydroxy- $\alpha$ -methyl cinnamoyl derivatives. Photoremovable protective groups are described in, for example, Patchornik, *J. Am. Chem. Soc.* (1970) 92:6333 and Amit et al., *J. Org. Chem.* (1974) 39:192, both of which are incorporated herein by reference.

Alternatively, the positive reactive group is activated for reaction with reagents in solution. For example, a 5-bromo-7-nitro indoline group, when bound to a carbonyl, undergoes reaction upon exposure to light at 420 nm. In a second alternative, the reactive group on the linker molecule is selected from a wide variety of negative light-reactive groups including a cinnamate group.

Alternatively, the reactive group is activated or deactivated by electron beam lithography, x-ray lithography, or any other radiation. Suitable reactive groups for electron beam lithography include sulfonyl. Other methods may be used including, for example, exposure to a current source. Other reactive groups and methods of activation may be used in light of this disclosure.

As shown in Fig. 1, the linking molecules are preferably exposed to, for example, light through a suitable mask 8 using photolithographic techniques of the type known in the semiconductor industry and described in, for example, Sze, *BLSI Technology*, McGraw-Hill (1983), and Mead et al., *Introduction to VLSI Systems*, Addison-Wesley (1980), which are incorporated herein by reference for all purposes. The light may be directed at either the surface containing the protective groups or at the back of the substrate, so long as the substrate is transparent to the wavelength of light needed for removal of the protective groups. In Fig. 1, light is directed at the surface of the substrate containing the protective groups. Fig. 1 illustrates the use of such masking techniques as they are applied to a positive reactive group so as to activate linking molecules and expose functional groups in areas 10a and 10b.

The mask 8 is, for example, a transparent support material selectively coated with a layer of opaque material.

Portions of the opaque material are removed, leaving opaque material in the precise pattern desired on the substrate surface. The mask is brought into close proximity with, imaged on, or brought directly into contact with the substrate surface as shown in Fig. 1. "Openings" in the mask correspond to locations on the substrate where it is desired to remove photoremovable protective groups from the substrate. Alignment may be performed using conventional alignment techniques in which alignment marks (not shown) are used to accurately overlay successive masks with previous patterning steps, or more sophisticated techniques may be used. For example, interferometric techniques such as the one described in Flanders *et al.*, "A New Interferometric Alignment Technique," *App. Phys. Lett.* (1977) 31:426-428, which is incorporated herein by reference, may be used.

To enhance contrast of light applied to the substrate, it can sometimes be desirable to provide contrast enhancement materials between the mask and the substrate. This contrast enhancement layer may comprise a molecule which is decomposed by light such as quinone diazid or a material which is transiently bleached at the wavelength of interest. Transient bleaching of materials will allow greater penetration where light is applied, thereby enhancing contrast. Alternatively, contrast enhancement may be provided by way of a cladded fiber optic bundle.

The light may be from a conventional incandescent source, a laser, a laser diode, or the like. If non-collimated sources of light are used it may be desirable to provide a thick- or multi-layered mask to prevent spreading of the light onto the substrate. It may, further, be desirable to utilize groups which are sensitive to different wavelengths to control synthesis. For example, by using groups which are sensitive to different wavelengths, it is possible to select branch positions in the synthesis of a polymer or eliminate certain masking steps. Several reactive groups along with their corresponding wavelengths for deprotection are provided in Table 1.

Table 1

Group	Approximate Deprotection Wavelength
Nitroveratryloxy carbonyl (NVOC)	UV (300-400 nm)
Nitrobenzyloxy carbonyl (NBOC)	UV (300-350 nm)
Dimethyl dimethoxybenzyloxy carbonyl	UV (280-300 nm)
5-Bromo-7-nitroindolyl	UV (420 nm)
<i>o</i> -Hydroxy- $\alpha$ -methyl cinnamoyl	UV (300-350 nm)
2-Oxymethylene anthraquinone	UV (350 nm)

While the present description refers primarily to the use of a mask to illuminate selected regions the substrate, other techniques may also be used. For example, the substrate may be translated under a modulated laser or diode light source. Such techniques are discussed in, for example, U.S. Patent No. 4,719,615 (Feyrer *et al.*), which is incorporated herein by reference. Alternatively, a laser galvanometric scanner is utilized. The synthesis may also take place on or in contact with a conventional liquid crystal (referred to herein as a "light valve") or fiber optic light sources. By appropriately modulating liquid crystals, light may be selectively controlled so as to permit light to contact selected regions of the substrate. Alternatively, synthesis may take place on the end of a series of optical fibers to which light is selectively applied. Other means of controlling the location of light exposure will be apparent to those of skill in the art.

The substrate may be irradiated either in contact or not in contact with a solution (not shown) and is, preferably, irradiated in contact with a solution. The solution can contain reagents to prevent the by-products formed by irradiation from interfering with synthesis of the polymer. Such by-products might include, for example, carbon dioxide, nitroso-carbonyl compounds, styrene derivatives indole derivatives, and products of their photochemical reactions. Alternatively, the solution may contain reagents used to match the index of refraction of the substrate. Reagents added to the solution may further include, for example, acidic or basic buffers, thiols, substituted hydrazines and hydroxylamines, reducing agents (e.g., NADH) or reagents known to react with a given functional group (e.g., aryl nitroso + glyoxylic acid  $\rightarrow$  aryl formhydroxamate + CO<sub>2</sub>).

Either concurrently with or after the irradiation step, the linker molecules are washed or otherwise contacted with a first monomer, illustrated by "A" in regions 12a and 12b in Fig. 2. The first monomer reacts with the activated functional groups of the linkage molecules which have been exposed to light. The first monomer, which is preferably an amino acid, is also provided with a photoprotective group. The photoprotective group on the monomer may be the same as or different than the protective group used in the linkage molecules, and may be selected from any of the above-described protective groups. The protective groups for the A monomer can be selected from the group NBOC and NVOC.

As shown in Fig. 3, the process of irradiating is thereafter repeated, with a mask repositioned so as to remove linkage protective groups and expose functional groups in regions 14a and 14b which are illustrated as being regions which were protected in the previous masking step. As an alternative to repositioning of the first mask, a second mask may be utilized. Alternatively, some steps may provide for illuminating a common region in successive steps. As shown

in Fig. 3, it may be desirable to provide separation between irradiated regions. For example, separation of about 1-5  $\mu\text{m}$  may be appropriate to account for alignment tolerances.

As shown in Fig. 4, the substrate is then exposed to a second protected monomer "B," producing B regions 16a and 16b. Thereafter, the substrate is again masked so as to remove the protective groups and expose reactive groups on A region 12a and B region 16b. The substrate is again exposed to monomer B, resulting in the formation of the structure shown in Fig. 6. The dimers B-A and B-B have been produced on the substrate.

A subsequent series of masking and contacting steps similar to those described above with A (not shown) provides the structure shown in Fig. 7. The process provides all possible dimers of B and A, i.e., B-A, A-B, A-A, and B-B.

The substrate, the area of synthesis, and the area for synthesis of each individual polymer could be of any size or shape. For example, squares, ellipsoids, rectangles, triangles, circles, or portions thereof, along with irregular geometric shapes, may be utilized. Duplicate synthesis areas may also be applied to a single substrate for purposes of redundancy.

The regions 12 to 16 on the substrate will have a surface area of less than about  $2.5 \times 10^{-3} \text{ cm}^2$ . In some embodiments the regions 12 and 16 have areas of less than about  $10^{-3} \text{ cm}^2$ ,  $10^{-4} \text{ cm}^2$ ,  $10^{-5} \text{ cm}^2$ ,  $10^{-6} \text{ cm}^2$ ,  $10^{-7} \text{ cm}^2$ ,  $10^{-8} \text{ cm}^2$ , or  $10^{-10} \text{ cm}^2$ . In a preferred embodiment, the regions 12 and 16 are between about  $10 \times 10 \mu\text{m}$  and  $500 \times 500 \mu\text{m}$ .

In some versions a single substrate supports more than about  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$ , or  $10^8$  different monomer sequences. Of course within a region of the substrate in which a monomer sequence is synthesized, it is preferred that the monomer sequence be substantially pure. In some embodiments, regions of the substrate contain polymer sequences which are at least about 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, or 99% pure.

According to some embodiments, several sequences are intentionally provided within a single region so as to provide an initial screening for biological activity, after which materials within regions exhibiting significant binding are further evaluated.

#### IV. Details of One Version of a Reactor System

Fig. 8A schematically illustrates a preferred reactor system 100 for synthesizing polymers on the prepared substrate. The reactor system includes a body 102 with a cavity 104 on a surface thereof. Preferably, the cavity 104 is between about 50 and  $1000 \mu\text{m}$  deep with a depth of about  $500 \mu\text{m}$  preferred.

The bottom of the cavity is preferably provided with an array of ridges 106 which extend both into the plane of the Figure and parallel to the plane of the Figure. The ridges are preferably about 50 to  $200 \mu\text{m}$  deep and spaced at about 2 to 3mm. The purpose of the ridges is to generate turbulent flow for better mixing. The bottom surface of the cavity is preferably light absorbing so as to prevent reflection of impinging light.

A substrate 112 is mounted above the cavity 104. The substrate is provided along its bottom surface 114 with a photoremovable protective group such as NVOC with or without an intervening linker molecule. The substrate is preferably transparent to a wide spectrum of light, but can be transparent only at a wavelength at which the protective group may be removed (such as UV in the case of NVOC). The substrate may be a conventional microscope glass slide or cover slip. The substrate is preferably as thin as possible, while still providing adequate physical support. Preferably, the substrate is less than about 1 mm thick, more preferably less than 0.5 mm thick, more preferably less than 0.1 mm thick, and most preferably less than 0.05 mm thick. In alternative preferred embodiments, the substrate is quartz or silicon.

The substrate and the body serve to seal the cavity except for an inlet port 108 and an outlet port 110. The body and the substrate may be mated for sealing with one or more gaskets. Preferably, the body is provided with two concentric gaskets and the intervening space is held at vacuum to ensure mating of the substrate is to the gaskets.

Fluid is pumped through the inlet port into the cavity by way of a pump 116 which may be, for example, a model no. B-120-S made by Eldex Laboratories. Selected fluids are circulated into the cavity by the pump, through the cavity, and out the outlet for recirculation or disposal. The reactor may be subjected to ultrasonic radiation and/or heated to aid in agitation in some embodiments.

Above the substrate 112, a lens 120 is provided which may be, for example, a 2" 100mm focal length fused silica lens. For the sake of a compact system, a reflective mirror 122 may be provided for directing light from a light source 124 onto the substrate. Light source 124 may be, for example, a Xe(Hg) light source manufactured by Oriol and having model no. 66024. A second lens 126 may be provided for the purpose of projecting a mask image onto the substrate in combination with lens 112. This form of lithography is referred to herein as projection printing. As will be apparent from this disclosure, proximity printing and the like may also be used.

Light from the light source is permitted to reach only selected locations on the substrate as a result of mask 128. Mask 128 may be, for example, a glass slide having etched chrome thereon. The mask 128 can be provided with a grid of transparent locations and opaque locations. Such masks may be manufactured by, for example, Photo sciences, Inc. Light passes freely through the transparent regions of the mask, but is reflected from or absorbed by other regions.

Therefore, only selected regions of the substrate are exposed to light.

As discussed above, light valves (LCD's) may be used as an alternative to conventional masks to selectively expose regions of the substrate. Alternatively, fiber optic faceplates such as those available from Schott Glass, Inc, may be used for the purpose of contrast enhancement of the mask or as the sole means of restricting the region to which light is applied. Such faceplates would be placed directly above or on the substrate in the reactor shown in Fig. 8A. In still further embodiments, flys-eye lenses, tapered fiber optic faceplates, or the like, may be used for contrast enhancement.

In order to provide for illumination of regions smaller than a wavelength of light, more elaborate techniques may be utilized. For example, light can be directed at the substrate by way of molecular microcrystals on the tip of, for example, micropipettes. Such devices are disclosed in Lieberman et al., "A Light Source Smaller Than the Optical Wavelength," *Science* (1990) 247:59-61, which is incorporated herein by reference for all purposes.

In operation, the substrate is placed on the cavity and sealed thereto. All operations in the process of preparing the substrate are carried out in a room lit primarily or entirely by light of a wavelength outside of the light range at which the protective group is removed. For example, in the case of NVOC, the room should be lit with a conventional dark room light which provides little or no UV light. All operations are preferably conducted at about room temperature.

A first, deprotection fluid (without a monomer) is circulated through the cavity. The solution preferably is of 5 mM sulfuric acid in dioxane solution which serves to keep exposed amino groups protonated and decreases their reactivity with photolysis by-products. Absorptive materials such as N,N-diethylamino 2,4-dinitrobenzene, for example, may be included in the deprotection fluid which serves to absorb light and prevent reflection and unwanted photolysis.

The slide is, thereafter, positioned in a light raypath from the mask such that first locations on the substrate are illuminated and, therefore, deprotected. The substrate is preferably illuminated for between about 1 and 15 minutes with a preferred illumination time of about 10 minutes at 10-20 mW/cm<sup>2</sup> with 365 nm light. The slides are neutralized (i.e., brought to a pH of about 7) after photolysis with, for example, a solution of diisopropylethylamine (DIEA) in methylene chloride for about 5 minutes.

The first monomer is then placed at the first locations on the substrate. After irradiation, the slide is removed, treated in bulk, and then reinstalled in the flow cell. Alternatively, a fluid containing the first monomer, preferably also protected by a protective group, is circulated through the cavity by way of pump 116. If, for example, it is desired to attach the amino acid Y to the substrate at the first locations, the amino acid Y (bearing a protective group on its  $\alpha$ -nitrogen), along with reagents used to render the monomer reactive, and/or a carrier, is circulated from a storage container 118, through the pump, through the cavity, and back to the inlet of the pump.

The monomer carrier solution is preferably formed by mixing of a first solution (referred to herein as solution "A") and a second solution (referred to herein as solution "B"). Table 2 provides an illustration of a mixture which may be used for solution A.

**Table 2**  
**Representative Monomer Carrier Solution "A"**

---

100 mg NVOC amino protected amino acid  
37 mg HOBT (1-Hydroxybenzotriazole)  
250  $\mu$ l DMF (Dimethylformamide)  
86  $\mu$ l DIEA (Diisopropylethylamine)

---

The composition of solution B is illustrated in Table 3. Solutions A and B are mixed and allowed to react at room temperature for about 8 minutes, then diluted with 2 ml of DMF, and 500  $\mu$ l are applied to the surface of the slide or the solution is circulated through the reactor system and allowed to react for about 2 hours at room temperature. The slide is then washed with DMF, methylene chloride and ethanol.

**Table 3**  
**Representative Monomer Carrier Solution "B"**

5

250  $\mu$ l DMF

10

111 mg BOP (Benzotriazolyl-n-oxy-tris(dimethylamino)  
 phosphoniumhexafluorophosphate)

15 As the solution containing the monomer to be attached is circulated through the cavity, the amino acid or other monomer will react at its carboxy terminus with amino groups on the regions of the substrate which have been deprotected. Of course, while the process is illustrated by way of circulation of the monomer through the cavity, the process could be practiced by way of removing the slide from the reactor and submersing it in an appropriate monomer solution.

After addition of the first monomer, the solution containing the first amino acid is then purged from the system. After circulation of a sufficient amount of the DMF/methylene chloride such that removal of the amino acid can be 20 assured (e.g., about 50x times the volume of the cavity and carrier lines), the mask or substrate is repositioned, or a new mask is utilized such that second regions on the substrate will be exposed to light and the light 124 is engaged for a second exposure. This will deprotect second regions on the substrate and the process is repeated until the desired polymer sequences have been synthesized.

25 The entire derivatized substrate is then exposed to a receptor of interest, preferably labeled with, for example, a fluorescent marker, by circulation of a solution or suspension of the receptor through the cavity or by contacting the surface of the slide in bulk. The receptor will preferentially bind to certain regions of the substrate which contain complementary sequences.

Antibodies are typically suspended in what is commonly referred to as "supercocktail," which may be, for example, a solution of about 1% BSA (bovine serum albumin), 0.5% Tween in PBS (phosphate buffered saline) buffer. The 30 antibodies are diluted into the supercocktail buffer to a final concentration of, for example, about 0.1 to 4  $\mu$ g/ml.

Fig. 8B illustrates an alternative preferred version of the reactor shown in Fig. 8A. According to this version, the mask 128 is placed directly in contact with the substrate. Preferably, the etched portion of the mask is placed face 35 down so as to reduce the effects of light dispersion. The imaging lenses 120 and 126 are not necessary because the mask is brought into close proximity with the substrate.

For purposes of increasing the signal-to-noise ratio of the technique, some embodiments of the process provide 40 for exposure of the substrate to a first labeled or unlabeled receptor followed by exposure of a labeled, second receptor (e.g., an antibody) which binds at multiple sites on the first receptor. If, for example, the first receptor is an antibody derived from a first species of an animal, the second receptor is an antibody derived from a second species directed to epitopes associated with the first species. In the case of a mouse antibody, for example, fluorescently labeled goat 45 antibody or antiserum which is antimouse may be used to bind at multiple sites on the mouse antibody, providing several times the fluorescence compared to the attachment of a single mouse antibody at each binding site. This process may be repeated again with additional antibodies (e.g., goat-mouse-goat, etc.) for further signal amplification.

Preferably, an ordered sequence of masks is utilized. It can be possible to use as few as a single mask to synthesize 50 all of the possible polymers of a given monomer set.

If, for example, it is desired to synthesize all 16 dinucleotides from four bases, a 1 cm square synthesis region is 45 divided conceptually into 16 boxes, each 0.25 cm wide. Denote the four monomer units by A, B, C, and D. The first reactions are carried out in four vertical columns, each 0.25 cm wide. The first mask exposes the left-most column of boxes, where A is coupled. The second mask exposes the next column, where B is coupled; followed by a third mask, 50 for the C column; and a final mask that exposes the right-most column, for D. The first, second, third, and fourth masks may be a single mask translated to different locations.

The process is repeated in the horizontal direction for the second unit of the dimer. This time, the masks allow exposure of horizontal rows, again 0.25 cm wide. A, B, C, and D are sequentially coupled using masks that expose 55 horizontal fourths of the reaction area. The resulting substrate contains all 16 dinucleotides of four bases.

The eight masks used to synthesize the dinucleotide are related to one another by translation or rotation. In fact, 60 one mask can be used in all eight steps if it is suitably rotated and translated. For example, in the example above, a mask with a single transparent region could be sequentially used to expose each of the vertical columns, translated 90°, and then sequentially used to allow exposure of the horizontal rows.

Tables 4 and 5 provide a simple computer program in Quick Basic for planning a masking program and a sample

output, respectively, for the synthesis of a polymer chain of three monomers ("residues") having three different monomers in the first level, four different monomers in the second level, and five different monomers in the third level in a striped pattern. The output of the program is the number of cells, the number of "stripes" (light regions) on each mask, and the amount of translation required for each exposure of the mask.

Table 4  
Mask Strategy Program

```

10  DEFINIT A-Z
    DIM b(20), w(20), l(500)
    F$ = "LPT1:"
    OPEN f$ FOR OUTPUT AS #1

15  jmax = 3          'Number of residues
    b(1) = 3: b(2) = 4: b(3) = 5      'Number of building blocks for res 1,2,3
    g = 1: lmax(1) = 1

    FOR j = 1 TO jmax: g = g * b(j): NEXT j

20  w(0) = 0: w(1) = g / b(1)

    PRINT #1, "MASK2.BAS ", DATE$, TIME$: PRINT #1,
    PRINT #1, USING "Number of residues-==": jmax
    FOR j = 1 TO jmax
25  PRINT #1, USING "      Residue ==      == building blocks": j: b(j)
    NEXT j
    PRINT #1, "
    PRINT #1, USING "Number of cells-====": g: PRINT #1,

30  FOR j = 2 TO jmax
    lmax(j) = lmax(j - 1) * b(j - 1)
    w(j) = w(j - 1) / b(j)
    NEXT j

    FOR j = 1 TO jmax
35  PRINT #1, USING "Mask for residue ==": j: PRINT #1,
    PRINT #1, USING "      Number of stripes-====": lmax(j)
    PRINT #1, USING "      Width of each stripe-====": w(j)
    FOR l = 1 TO lmax(j)
    a = 1 + (l - 1) * w(j - 1)
40  ae = a + w(j) - 1
    PRINT #1, USING "      Stripe == begins at location === and ends at ===": l: a: ae
    NEXT l
    PRINT #1,
    PRINT #1, USING "      For each of == building blocks, translate mask by ==
45  cell(s)": b(j): w(j).
    PRINT #1, : PRINT #1, : PRINT #1,
    NEXT j

```

50 © Copyright 1990, Affymax N.V.

55

Table 5Masking Strategy Output


---

5  
 Number of residues- 3  
     Residue 1       3 building blocks  
     Residue 2       4 building blocks  
     Residue 3       5 building blocks

10  
 Number of cells- 60  
 Mask for residue 1

15  
     Number of stripes- 1  
     Width of each stripe- 20  
     Stripe 1 begins at location 1 and ends at 20  
  
     For each of 3 building blocks, translate mask by 20 cell(s)

20  
 Mask for residue 2

25  
     Number of stripes- 3  
     Width of each stripe- 5  
     Stripe 1 begins at location 1 and ends at 5  
     Stripe 2 begins at location 21 and ends at 25  
     Stripe 3 begins at location 41 and ends at 45

30  
     For each of 4 building blocks, translate mask by 5 cell(s)

Mask for residue 3

35  
     Number of stripes- 12  
     Width of each stripe- 1  
     Stripe 1 begins at location 1 and ends at 1  
     Stripe 2 begins at location 6 and ends at 6  
     Stripe 3 begins at location 11 and ends at 11  
     Stripe 4 begins at location 16 and ends at 16  
     Stripe 5 begins at location 21 and ends at 21  
     Stripe 6 begins at location 26 and ends at 26  
     Stripe 7 begins at location 31 and ends at 31  
     Stripe 8 begins at location 36 and ends at 36  
     Stripe 9 begins at location 41 and ends at 41  
     Stripe 10 begins at location 46 and ends at 46  
     Stripe 11 begins at location 51 and ends at 51  
     Stripe 12 begins at location 56 and ends at 56

45  
     For each of 5 building blocks, translate mask by 1 cell(s)

---

## V. Details of One Version of A Fluorescent Detection Device

Fig. 9 illustrates a fluorescent detection device for detecting fluorescently labeled receptors on a substrate. A substrate 112 is placed on an x/y translation table 202. Preferably, the x/y translation table is a model no. PM500-AI manufactured by Newport Corporation. The x/y translation table is connected to and controlled by an appropriately programmed digital computer 204 which may be, for example, an appropriately programmed IBM PC/AT or AT compatible computer. Of course, other computer systems, special purpose hardware, or the like could readily be substituted for the AT computer used herein for illustration. Computer software for the translation and data collection functions described herein can be provided based on commercially available software including, for example, "Lab Windows" licensed by National Instruments, which is incorporated herein by reference for all purposes.

The substrate and x/y translation table are placed under a microscope 206 which includes one or more objectives 208. Light (about 488 nm) from a laser 210, which in some embodiments is a model no. 2020-05 argon ion laser manufactured by Spectraphysics, is directed at the substrate by a dichroic mirror 207 which passes greater than about 520 nm light but reflects 488 nm light. Dichroic mirror 207 may be, for example, a model no. FT510 manufactured by Carl Zeiss. Light reflected from the mirror then enters the microscope 206 which may be, for example, a model no. Axioscop 20 manufactured by Carl Zeiss. Fluorescein-marked materials on the substrate will fluoresce >488 nm light, and the fluoresced light will be collected by the microscope and passed through the mirror. The fluorescent light from the substrate is then directed through a wavelength filter 209 and, thereafter through an aperture plate 211. Wavelength filter 209 may be, for example, a model no. OG530 manufactured by Melles Griot and aperture plate 211 may be, for example, a model no. 477352/477380 manufactured by Carl Zeiss.

The fluoresced light then enters a photomultiplier tube 212 which can be is a model no. R943-02 manufactured by Hamamatsu, the signal is amplified in preamplifier 214 and photons is recorded as a function of the location in the computer 204. Pre-Amp 214 may be, for example, a model no. SR440 manufactured by Stanford Research Systems and photon counter 216 may be a model no. SR400 manufactured by Stanford Research Systems. The substrate is then moved to a subsequent location and the process is repeated. In preferred embodiments the data are acquired every 1 to 100  $\mu\text{m}$  with a data collection diameter of about 0.8 to 10  $\mu\text{m}$  preferred. Where there is sufficiently high fluorescence, a CCD detector with broadfield illumination is utilized.

By counting the number of photons generated in a given area in response to the laser, it is possible to determine where fluorescent marked molecules are located on the substrate. Consequently, for a slide which has a matrix of polypeptides, for example, synthesized on the surface thereof, it is possible to determine which of the polypeptides is complementary to a fluorescently marked receptor.

Preferably, the intensity and duration of the light applied to the substrate is controlled by varying the laser power and scan stage rate for improved signal-to-noise ratio by maximizing fluorescence emission and minimizing background noise.

While the detection apparatus has been illustrated primarily herein with regard to the detection of marked receptors, the invention will find application in other areas. For example, the detection apparatus disclosed herein could be used in the fields of catalysis, DNA or protein gel scanning, and the like.

## VI. Determination of Relative Binding Strength of Receptors

The signal-to-noise ratio of the present invention is sufficiently high that not only can the presence or absence of a receptor on a ligand be detected, but also the relative binding affinity of receptors to a variety of sequences can be determined.

In practice it is found that a receptor will bind to several peptide sequences in an array, but will bind much more strongly to some sequences than others. Strong binding affinity will be evidenced herein by a strong fluorescent or radiographic signal since many receptor molecules will bind in a region of a strongly bound ligand. Conversely, a weak binding affinity will be evidenced by a weak fluorescent or radiographic signal due to the relatively small number of receptor molecules which bind in a particular region of a substrate having a ligand with a weak binding affinity for the receptor. Consequently, it becomes possible to determine relative binding avidity (or affinity in the case of univalent interactions) of a ligand herein by way of the intensity of a fluorescent or radiographic signal in a region containing that ligand.

Semiquantitative data on affinities might also be obtained by varying washing conditions and concentrations of the receptor. This would be done by comparison to known ligand receptor pairs, for example.

## VII. Examples

The following examples are provided to illustrate the efficacy of the invention herein. All operations were conducted at about ambient temperatures and pressures unless indicated to the contrary.



## A. Slide Preparation

Before attachment of reactive groups it is preferred to clean the substrate which is, in a preferred embodiment a glass substrate such as a microscope slide or cover slip. According to one embodiment the slide is soaked in an alkaline bath consisting of, for example, 1 liter of 95% ethanol with 120 ml of water and 120 grams of sodium hydroxide for 12 hours. The slides are then washed under running water and allowed to air dry, and rinsed once with a solution of 95% ethanol.

The slides are then aminated with, for example, aminopropyltriethoxysilane for the purpose of attaching amino groups to the glass surface on linker molecules, although any omega functionalized silane could also be used for this purpose. In one embodiment 0.1% aminopropyltriethoxysilane is utilized, although solutions with concentrations from 10<sup>-7</sup>% to 10% may be used, with about 10<sup>-3</sup>% to 2% preferred. A 0.1% mixture is prepared by adding to 100 ml of a 95% ethanol/5% water mixture, 100 microliters (μl) of aminopropyltriethoxysilane. The mixture is agitated at about ambient temperature on a rotary shaker for about 5 minutes. 500 μl of this mixture is then applied to the surface of one side of each cleaned slide. After 4 minutes, the slides are decanted of this solution and rinsed three times by dipping in, for example, 100% ethanol.

After the plates dry, they are placed in a 110-120°C vacuum oven for about 20 minutes, and then allowed to cure at room temperature for about 12 hours in an argon environment. The slides are then dipped into DMF (dimethylformamide) solution, followed by a thorough washing with methylene chloride.

The aminated surface of the slide is then exposed to about 500 μl of, for example, a 30 millimolar (mM) solution of NVOC-GABA (gamma amino butyric acid) NHS (N-hydroxysuccinimide) in DMF for attachment of a NVOC-GABA to each of the amino groups.

The surface is washed with, for example, DMF, methylene chloride, and ethanol.

Any unreacted aminopropyl silane on the surface—that is, those amino groups which have not had the NVOC-GABA attached—are now capped with acetyl groups (to prevent further reaction) by exposure to a 1:3 mixture of acetic anhydride in pyridine for 1 hour. Other materials which may perform this residual capping function include trifluoroacetic anhydride, formicacetic anhydride, or other reactive acylating agents. Finally, the slides are washed again with DMF, methylene chloride, and ethanol.

## B. Synthesis of Eight Trimers of "A" and "B"

Fig. 10 illustrates a possible synthesis of the eight trimers of the two-monomer set: gly, phe (represented by "A" and "B," respectively). A glass slide bearing silane groups terminating in 6-nitro-veratryloxycarboxamide (NVOC-NH) residues is prepared as a substrate. Active esters (pentafluorophenyl, OBt, etc.) of gly and phe protected at the amino group with NVOC are prepared as reagents. While not pertinent to this example, if side chain protecting groups are required for the monomer set, these must not be photoreactive at the wavelength of light used to protect the primary chain.

For a monomer set of size  $n$ ,  $n \times \ell$  cycles are required to synthesize all possible sequences of length  $\ell$ . A cycle consists of:

1. Irradiation through an appropriate mask to expose the amino groups at the sites where the next residue is to be added, with appropriate washes to remove the by-products of the deprotection.
2. Addition of a single activated and protected (with the same photochemically-removable group) monomer, which will react only at the sites addressed in step 1, with appropriate washes to remove the excess reagent from the surface.

The above cycle is repeated for each member of the monomer set until each location on the surface has been extended by one residue in one embodiment. In other embodiments, several residues are sequentially added at one location before moving on to the next location. Cycle times will generally be limited by the coupling reaction rate, now as short as 20 min in automated peptide synthesizers. This step is optionally followed by addition of a protecting group to stabilize the array for later testing. For some types of polymers (e.g., peptides), a final deprotection of the entire surface (removal of photoprotective side chain groups) may be required.

More particularly, as shown in Fig. 10A, the glass is provided with regions 22, 24, 26, 28, 30, 32, 34, and 36. Regions 30, 32, 34, and 36 are masked, as shown in Fig. 10B and the glass is irradiated and exposed to a reagent containing "A" (e.g., gly), with the resulting structure shown in Fig. 10C. Thereafter, regions 22, 24, 26, and 28 are masked, the glass is irradiated (as shown in Fig. 10D) and exposed to a reagent containing "B" (e.g., phe), with the resulting structure shown in Fig. 10E. The process proceeds, consecutively masking and exposing the sections as shown until the structure shown in Fig. 10M is obtained. The glass is irradiated and the terminal groups are, optionally, capped by acetylation. As shown, all possible trimers of gly/phe are obtained.

## EP 0 619 321 B1

In this example, no side chain protective group removal is necessary. If it is desired, side chain deprotection may be accomplished by treatment with ethanedithiol and trifluoroacetic acid.

In general, the number of steps needed to obtain a particular polymer chain is defined by:

$$n \times \ell \quad (1)$$

where:

$n$  = the number of monomers in the basis set of monomers, and  
 $\ell$  = the number of monomer units in a polymer chain.

Conversely, the synthesized number of sequences of length  $\ell$  will be:

$$n^\ell \quad (2)$$

Of course, greater diversity is obtained by using masking strategies which will also include the synthesis of polymers having a length of less than  $\ell$ . If, in the extreme case, all polymers having a length less than or equal to  $\ell$  are synthesized, the number of polymers synthesized will be:

$$n^\ell + n^{\ell-1} + \dots + n^1 \quad (3)$$

The maximum number of lithographic steps needed will generally be  $n$  for each "layer" of monomers, i.e., the total number of masks (and, therefore, the number of lithographic steps) needed will be  $n \times \ell$ . The size of the transparent mask regions will vary in accordance with the area of the substrate available for synthesis and the number of sequences to be formed. In general, the size of the synthesis areas will be:

$$\text{size of synthesis areas} = (A)/(S)$$

where:

$A$  is the total area available for synthesis; and  
 $S$  is the number of sequences desired in the area.

It will be appreciated by those of skill in the art that the above method could readily be used to simultaneously produce thousands or millions of oligomers on a substrate using the photolithographic techniques disclosed herein. Consequently, the method results in the ability to practically test large numbers of, for example, di, tri, tetra, penta, hexa, hepta, octapeptides, dodecapeptides, or larger polypeptides (or correspondingly, polynucleotides).

The above example has illustrated the method by way of a manual example. It will of course be appreciated that automated or semi-automated methods could be used. The substrate would be mounted in a flow cell for automated addition and removal of reagents, to minimize the volume of reagents needed, and to more carefully control reaction conditions. Successive masks could be applied manually or automatically.

### C. Synthesis of a Dimer of an Aminopropyl Group and a Fluorescent Group

In synthesizing the dimer of an aminopropyl group and a fluorescent group, a functionalized durapore membrane was used as a substrate. The durapore membrane was a polyvinylidene difluoride with aminopropyl groups. The aminopropyl groups were protected with the DDZ group by reaction of the carbonyl chloride with the amino groups, a reaction readily known to those of skill in the art. The surface bearing these groups was placed in a solution of THF and contacted with a mask bearing a checkerboard pattern of 1 mm opaque and transparent regions. The mask was exposed to ultraviolet light having a wavelength down to at least about 280 nm for about 5 minutes at ambient temperature, although a wide range of exposure times and temperatures may be appropriate in various embodiments of the invention. For example, in one embodiment, an exposure time of between about 1 and 5000 seconds may be used at process temperatures of between -70 and +50°C.

In one preferred embodiment, exposure times of between about 1 and 500 seconds at about ambient pressure are used. In some preferred embodiments, pressure above ambient is used to prevent evaporation.

The surface of the membrane was then washed for about 1 hour with a fluorescent label which included an active ester bound to a chelate of a lanthanide. Wash times will vary over a wide range of values from about a few minutes to a few hours. These materials fluoresce in the red and the green visible region. After the reaction with the active ester in the fluorophore was complete, the locations in which the fluorophore was bound could be visualized by exposing them to ultraviolet light and observing the red and the green fluorescence. It was observed that the derivatized regions of the substrate closely corresponded to the original pattern of the mask.

#### D. Demonstration of Signal Capability

Signal detection capability was demonstrated using a low-level standard fluorescent bead kit manufactured by Flow Cytometry Standards and having model no. 824. This kit includes 5.8  $\mu\text{m}$  diameter beads, each impregnated with a known number of fluorescein molecules.

One of the beads was placed in the illumination field on the scan stage as shown in Fig. 9 in a field of a laser spot which was initially shuttered. After being positioned in the illumination field, the photon detection equipment was turned on. The laser beam was unblocked and it interacted with the particle bead, which then fluoresced. Fluorescence curves of beads impregnated with 7,000 and 29,000 fluorescein molecules, are shown in Figs. 11A and 11B, respectively. On each curve, traces for beads without fluorescein molecules are also shown. These experiments were performed with 488 nm excitation, with 100  $\mu\text{W}$  of laser power. The light was focused through a 40 power 0.75 NA objective.

The fluorescence intensity in all cases started off at a high value and then decreased exponentially. The fall-off in intensity is due to photobleaching of the fluorescein molecules. The traces of beads without fluorescein molecules are used for background subtraction. The difference in the initial exponential decay between labeled and nonlabeled beads is integrated to give the total number of photon counts, and this number is related to the number of molecules per bead. Therefore, it is possible to deduce the number of photons per fluorescein molecule that can be detected. For the curves illustrated in Fig. 11, this calculation indicates the radiation of about 40 to 50 photons per fluorescein molecule are detected.

#### E. Determination of the Number of Molecules Per Unit Area

Aminopropylated glass microscope slides prepared according to the methods discussed above were utilized in order to establish the density of labeling of the slides. The free amino termini of the slides were reacted with FITC (fluorescein isothiocyanate) which forms a covalent linkage with the amino group. The slide is then scanned to count the number of fluorescent photons generated in a region which, using the estimated 40-50 photons per fluorescent molecule, enables the calculation of the number of molecules which are on the surface per unit area.

A slide with aminopropyl silane on its surface was immersed in a 1 mM solution of FITC in DMF for 1 hour at about ambient temperature. After reaction, the slide was washed twice with DMF and then washed with ethanol, water, and then ethanol again. It was then dried and stored in the dark until it was ready to be examined.

Through the use of curves similar to those shown in Fig. 11, and by integrating the fluorescent counts under the exponentially decaying signal, the number of free amino groups on the surface after derivitization was determined. It was determined that slides with labeling densities of 1 fluorescein per  $10^3 \times 10^3$  to  $\sim 2 \times 2$  nm could be reproducibly made as the concentration of aminopropyltriethoxysilane varied from  $10^{-5}\%$  to  $10^{-1}\%$ .

#### F. Removal of NVOC and Attachment of A Fluorescent Marker

NVOC-GABA groups were attached as described above. The entire surface of one slide was exposed to light so as to expose a free amino group at the end of the gamma amino butyric acid. This slide, and a duplicate which was not exposed, were then exposed to fluorescein isothiocyanate (FITC).

Fig. 12A illustrates the slide which was not exposed to light, but which was exposed to FITC. The units of the x axis are time and the units of the y axis are counts. The trace contains a certain amount of background fluorescence. The duplicate slide was exposed to 350 nm broadband illumination for about 1 minute ( $12 \text{ mW/cm}^2$ ,  $\sim 350 \text{ nm}$  illumination), washed and reacted with FITC. The fluorescence curves for this slide are shown in Fig. 12B. A large increase in the level of fluorescence is observed, which indicates photolysis has exposed a number of amino groups on the surface of the slides for attachment of a fluorescent marker.

#### G. Use of a Mask in Removal of NVOC

The next experiment was performed with a 0.1% aminopropylated slide. Light from a Hg-Xe arc lamp was imaged

onto the substrate through a laser-ablated chrome-on-glass mask in direct contact with the substrate.

This slide was illuminated for approximately 5 minutes, with 12 mW of 350 nm broadband light and then reacted with the 1 mM FITC solution. It was put on the laser detection scanning stage and a graph was plotted as a two-dimensional representation of position color-coded for fluorescence intensity. The experiment was repeated a number of times through various masks. The fluorescence patterns for a 100x100  $\mu\text{m}$  mask, a 50  $\mu\text{m}$  mask, a 20  $\mu\text{m}$  mask, and a 10  $\mu\text{m}$  mask indicate that the mask pattern is distinct down to at least about 10  $\mu\text{m}$  squares using this lithographic technique.

#### H. Attachment of YGGFL and Subsequent Exposure to Herz Antibody and Goat Antimouse

In order to establish that receptors to a particular polypeptide sequence would bind to a surface-bound peptide and be detected, Leu enkephalin was coupled to the surface and recognized by an antibody. A slide was derivatized with 0.1% amino propyl-triethoxysilane and protected with NVOC. A 500  $\mu\text{m}$  checkerboard mask was used to expose the slide in a flow cell using backside contact printing. The Leu enkephalin sequence ( $\text{H}_2\text{N}$ -tyrosine, glycine, glycine, phenylalanine, leucine- $\text{CO}_2\text{H}$ , otherwise referred to herein as YGGFL) was attached via its carboxy end to the exposed amino groups on the surface of the slide. The peptide was added in DMF solution with the BOP/HOBT/DIEA coupling reagents and recirculated through the flow cell for 2 hours at room temperature.

A first antibody, known as the Herz antibody, was applied to the surface of the slide for 45 minutes at 2  $\mu\text{g}/\text{ml}$  in a supercocktail (containing 1% BSA and 1% ovalbumin also in this case). A second antibody, goat anti-mouse fluorescein conjugate, was then added at 2  $\mu\text{g}/\text{ml}$  in the supercocktail buffer, and allowed to incubate for 2 hours.

The results of this experiment were plotted as fluorescence intensity as a function of position. This image was taken at 10  $\mu\text{m}$  steps and showed that not only can deprotection be carried out in a well defined pattern, but also that (1) the method provided for successful coupling of peptides to the surface of the substrate, (2) the surface of a bound peptide was available for binding with an antibody, and (3) that the detection apparatus capabilities were sufficient to detect binding of a receptor.

#### I. Monomer-by-Monomer Formation of YGGFL and Subsequent Exposure to Labeled Antibody

Monomer-by-monomer synthesis of YGGFL and GGFL in alternate squares was performed on a slide in a checkerboard pattern and the resulting slide was exposed to the Herz antibody. This experiment is illustrated in Figs. 13A and 13B.

In Fig. 13A, a slide is shown which is derivatized with the aminopropyl group, protected in this case with t-BOC (t-butoxycarbonyl). The slide was treated with TFA to remove the t-BOC protecting group. E-aminocaproic acid, which was t-BOC protected at its amino group, was then coupled onto the aminopropyl groups. The aminocaproic acid serves as a spacer between the aminopropyl group and the peptide to be synthesized. The amino end of the spacer was deprotected and coupled to NVOC-leucine. The entire slide was then illuminated with 12 mW of 325 nm broadband illumination. The slide was then coupled with NVOC-phenylalanine and washed. The entire slide was again illuminated, then coupled to NVOC-glycine and washed. The slide was again illuminated and coupled to NVOC-glycine to form the sequence shown in the last portion of Fig. 13A.

As shown in Fig. 13B, alternating regions of the slide were then illuminated using a projection print using a 500x500  $\mu\text{m}$  checkerboard mask; thus, the amino group of glycine was exposed only in the lighted areas. When the next coupling chemistry step was carried out, NVOC-tyrosine was added, and it coupled only at those spots which had received illumination. The entire slide was then illuminated to remove all the NVOC groups, leaving a checkerboard of YGGFL in the lighted areas and in the other areas, GGFL. The Herz antibody (which recognizes the YGGFL, but not GGFL) was then added, followed by goat anti-mouse fluorescein conjugate.

The resulting fluorescence scan showed dark areas containing the tetrapeptide GGFL, which is not recognized by the Herz antibody (and thus there is no binding of the goat anti-mouse antibody with fluorescein conjugate), and red areas in which YGGFL was present. The YGGFL pentapeptide is recognized by the Herz antibody and, therefore, there is antibody in the lighted regions for the fluorescein-conjugated goat anti-mouse to recognize.

Similar patterns for a 50  $\mu\text{m}$  mask used in direct contact ("proximity print") with the substrate provided a pattern which was more distinct and the corners of the checkerboard pattern were touching as a result of the mask being placed in direct contact with the substrate (which reflects the increase in resolution using this technique).

#### J. Monomer-by-Monomer Synthesis of YGGFL and PGGFL

A synthesis using a 50  $\mu\text{m}$  checkerboard mask similar to that shown in Fig. 13 was conducted. However, P was added to the GGFL sites on the substrate through an additional coupling step. P was added by exposing protected GGFL to light through a mask, and subsequent exposure to P in the manner set forth above. Therefore, half of the

regions on the substrate contained YGGFL and the remaining half contained PGGFL.

The fluorescence plot for this experiment showed the regions are again readily discernable between those in which binding did and did not occur. This experiment demonstrated that antibodies are able to recognize a specific sequence and that the recognition is not length-dependent.

#### K. Monomer-by-Monomer Synthesis of YGGFL and YPGGFL

In order to further demonstrate the operability of the invention, a 50  $\mu$ m checkerboard pattern of alternating YGGFL and YPGGFL was synthesized on a substrate using techniques like those set forth above. The resulting fluorescence plot showed that the antibody was clearly able to recognize the YGGFL sequence and did not bind significantly at the YPGGFL regions.

#### L. Synthesis of an Array of Sixteen Different Amino Acid Sequences and Estimation of Relative Binding Affinity to Herz Antibody

Using techniques similar to those set forth above, an array of 16 different amino acid sequences (replicated four times) was synthesized on each of two glass substrates. The sequences were synthesized by attaching the sequence NVOC-GFL across the entire surface of the slides. Using a series of masks, two layers of amino acids were then selectively applied to the substrate. Each region had dimensions of 0.25 cm x 0.0625 cm. The first slide contained amino acid sequences containing only L amino acids while the second slide contained selected D amino acids. Figs. 14A and 14B illustrate a map of the various regions on the first and second slides, respectively. The patterns shown in Figs. 14A and 14B were duplicated four times on each slide. The slides were then exposed to the Herz antibody and fluorescein-labeled goat anti-mouse.

A fluorescence plot of the first slide, which contained only L amino acids showed red areas (indicating strong binding, i.e., 149,000 counts or more) and black areas (indicating little or no binding of the Herz antibody, i.e., 20,000 counts or less). The sequence YGGFL was clearly most strongly recognized. The sequences YAGFL and YSGFL also exhibited strong recognition of the antibody. By contrast, most of the remaining sequences showed little or no binding. The four duplicate portions of the slide were extremely consistent in the amount of binding shown therein.

A fluorescence plot of the D amino acid slide indicated that strongest binding was exhibited by the YGGFL sequence. Significant binding was also detected to YaGFL, YsGFL, and YpGFL. The remaining sequences showed less binding with the antibody. Low binding efficiency of the sequence yGGFL was observed.

Table 6 lists the various sequences tested in order of relative fluorescence, which provides information regarding relative binding affinity.

Table 6

Apparent Binding to Herz Ab	
L-a.a. Set	D-a.a. Set
YGGFL	YGGFL
YAGFL	YaGFL
YSGFL	YsGFL
LGGFL	YpGFL
FGGFL	fGGFL
YPGFL	yGGFL
LAGFL	laGFL
FAGFL	wGGFL
WGGFL	yaGFL
	fpGFL
	waGFL

#### VIII. Illustrative Alternative Approach

According to an alternative approach, the methods provide for attaching to the surface a caged binding member which in its caged form has a relatively low affinity for other potentially binding species, such as receptors and specific binding substances. Such techniques are more fully described in copending application Serial No. 404,920, filed September 8, 1989, and incorporated herein by reference for all purposes.

According to this alternative methods are provided for forming predefined regions on a surface of a solid support, wherein the predefined regions are capable of immobilizing receptors. The methods make use of caged binding members attached to the surface to enable selective activation of the predefined regions. The caged binding members are liberated to act as binding members ultimately capable of binding receptors upon selective activation of the predefined regions. The activated binding members are then used to immobilize specific molecules such as receptors on the predefined region of the surface. The above procedure is repeated at the same or different sites on the surface so as to provide a surface prepared with a plurality of regions on the surface containing, for example, the same or different receptors. When receptors immobilized in this way have a differential affinity for one or more ligands, screenings and assays for the ligands can be conducted in the regions of the surface containing the receptors.

The alternative approach may make use of novel caged binding members attached to the substrate. Caged (un-activated) members have a relatively low affinity for receptors of substances that specifically bind to uncaged binding members when compared with the corresponding affinities of activated binding members. Thus, the binding members are protected from reaction until a suitable source of energy is applied to the regions of the surface desired to be activated. Upon application of a suitable energy source, the caging groups labilize, thereby presenting the activated binding member. A typical energy source will be light.

Once the binding members on the surface are activated they may be attached to a receptor. The receptor chosen may be a monoclonal antibody, a nucleic acid sequence, a drug receptor, etc. The receptor will usually, though not always, be prepared so as to permit attaching it, directly or indirectly, to a binding member. For example, a specific binding substance having a strong binding affinity for the binding member and a strong affinity for the receptor or a conjugate of the receptor may be used to act as a bridge between binding members and receptors if desired. The method uses a receptor prepared such that the receptor retains its activity toward a particular ligand.

Preferably, the caged binding member attached to the solid substrate will be a photoactivatable biotin complex, i. e., a biotin molecule that has been chemically modified with photoactivatable protecting groups so that it has a significantly reduced binding affinity for avidin or avidin analogs than does natural biotin. Preferably, the protecting groups localized in a predefined region of the surface will be removed upon application of a suitable source of radiation to give binding members, that are biotin or a functionally analogous compound having substantially the same binding affinity for avidin or avidin analogs as does biotin.

In another preferred approach, avidin or an avidin analog is incubated with activated binding members on the surface until the avidin binds strongly to the binding members. The avidin so immobilized on predefined regions of the surface can then be incubated with a desired receptor or conjugate of a desired receptor. The receptor will preferably be biotinylated, e.g., a biotinylated antibody, when avidin is immobilized on the predefined regions of the surface. Alternatively, an avidin/biotinylated receptor complex can be used, which has been previously prepared, to activated binding members on the surface.

#### IX. Conclusion

It is to be understood that the above description is intended to be illustrative and not restrictive. Many embodiments will be apparent to those of skill in the art upon reviewing the above description. By way of example, the invention has been described primarily with reference to the use of photoremovable protective groups, but it will be readily recognized by those of skill in the art that sources of radiation other than light could also be used. For example, it may be desirable to use protective groups which are sensitive to electron beam irradiation, x-ray irradiation, in combination with electron beam lithograph, or x-ray lithography techniques. Alternatively, the group could be removed by exposure to an electric current.

#### Claims

1. A method of investigating by receptor/ligand binding a polynucleotide or amino acid sequence by the use of a substrate with a surface, said surface comprising at least  $10^3$  predefined regions, said predefined regions containing different nucleotide or amino acid sequences thereon, said predefined regions each occupying an area of less than about  $2.5 \times 10^{-3} \text{ cm}^2$ , which method comprises labelling said sequence being investigated and identifying which of said different sequences binds with said sequence being analysed.
2. A method as claimed in claim 1, wherein said area is less than  $10,000 \mu\text{m}^2$ , optionally less than about  $100 \mu\text{m}^2$ .
3. A method as claimed in claim 1 or claim 2, wherein there are  $10^4$  or more different monomer sequences and predefined regions, optionally  $10^5$  or more, on said surface.

4. A method as claimed in any one of claims 1 to 3, wherein said label is a fluorescent label, a radioactive label, or a labelled antibody.
5. A method as claimed in any one of claims 1 to 4, wherein the monomer sequences in the predefined regions where binding is detected are used to determine all or part of a sequence complementary to the sequence being analysed.
6. A method as claimed in any one of claims 1 to 5, wherein the monomer sequences are from 2 to 20 monomers long.
7. Apparatus for investigating by receptor/ligand binding a polynucleotide or amino acid sequence comprising a substrate with a surface, said surface comprising at least  $10^3$  predefined regions, said predefined regions containing different polynucleotide or amino acid sequences thereon, said predefined regions each occupying an area of less than about  $2.5 \times 10^{-3} \text{ cm}^2$ .
8. Apparatus as claimed in claim 7 and further defined by the specific feature(s) of any one or more of claims 2, 3 or 6.
9. Apparatus as claimed in claim 7 or claim 8, wherein said substrate is less than 0.5 mm thick, optionally less than 0.05 mm thick.
10. Apparatus as claimed in any one of claims 7 to 9, wherein said predefined regions have a surface area of less than about  $10^{-6} \text{ cm}^2$ .
11. Apparatus as claimed in any one of claims 7 to 10, wherein each polynucleotide or amino acid monomer sequence is at least 50%, preferably at least 90% pure within a predefined region.
12. The use of an apparatus as defined in any one of claims 7 to 11 in investigating by receptor/ligand binding a polynucleotide or amino acid sequence.

#### Patentansprüche

1. Verfahren zur Untersuchung einer Polynukleotid- oder einer Aminosäuresequenz mittels Rezeptor/Liganden-Bindung unter Verwendung eines Substrats mit einer Oberfläche, wobei die Oberfläche mindestens  $10^3$  vorbestimmte Bereiche umfaßt, die vorbestimmten Bereiche unterschiedliche Nukleotid- oder Aminosäuresequenzen enthalten, die vorbestimmten Bereiche jeweils einen Bereich von weniger als etwa  $2,5 \times 10^{-3} \text{ cm}^2$  besitzen, wobei das Verfahren die Markierung der untersuchten Sequenz und das Identifizieren umfaßt, welche der unterschiedlichen Sequenzen mit der zu analysierenden Sequenz binden.
2. Verfahren nach Anspruch 1, wobei der Bereich weniger als  $10.000 \mu\text{m}^2$ , gegebenenfalls weniger als etwa  $100 \mu\text{m}^2$  beträgt.
3. Verfahren nach Anspruch 1 oder 2, wobei sich auf der Oberfläche  $10^4$  oder mehr, gegebenenfalls  $10^5$  oder mehr, unterschiedliche Monomer-Sequenzen und vorbestimmte Bereiche befinden.
4. Verfahren nach einem der Ansprüche 1 bis 3, wobei die Markierung eine fluoreszierende Markierung, eine radioaktive Markierung oder ein markierter Antikörper ist.
5. Verfahren nach einem der Ansprüche 1 bis 4, wobei die Monomer-Sequenzen in den vorbestimmten Bereichen, in denen die Bindung festgestellt wird, zur Bestimmung sämtlicher oder eines Teils einer Sequenz verwendet wird, die komplementär zu der zu analysierenden Sequenz ist.
6. Verfahren nach einem der Ansprüche 1 bis 5, wobei die Monomer-Sequenzen 2 bis 20 Monomere lang sind.
7. Vorrichtung zur Untersuchung einer Polynukleotid- oder einer Aminosäuresequenz mittels Rezeptor/Liganden-Bindung, umfassend ein Substrat mit einer Oberfläche, wobei die Oberfläche mindestens  $10^3$  vorbestimmte Bereiche umfaßt, die vorbestimmten Bereiche unterschiedliche Polynukleotid- oder Aminosäuresequenzen enthalten, und die vorbestimmten Bereiche jeweils einen Bereich von weniger als etwa  $2,5 \times 10^{-3} \text{ cm}^2$  besitzen.
8. Vorrichtung nach Anspruch 7, weiter gekennzeichnet durch die speziellen Merkmale nach einem oder mehreren

der Ansprüche 2, 3 oder 6.

9. Vorrichtung nach Anspruch 7 oder 8, wobei das Substrat weniger als 0,5 mm, gegebenenfalls weniger als 0,05 mm dick ist.

10. Vorrichtung nach einem der Ansprüche 7 bis 9, wobei die vorbestimmten Bereiche eine Oberfläche von weniger als etwa  $10^{-5}$  cm<sup>2</sup> haben.

11. Vorrichtung nach einem der Ansprüche 7 bis 10, wobei jede Polynukleotid- oder Aminosäure-Monomer-Sequenz mindestens 50%, vorzugsweise mindestens 90% rein innerhalb eines vorbestimmten Bereiches ist.

12. Verwendung einer Vorrichtung nach einem der Ansprüche 7 bis 11 zur Untersuchung einer Polynukleotid- oder einer Aminosäuresequenz mittels Rezeptor/Liganden-Bindung.

# Revendications

1. Procédé d'investigation par liaison récepteur/ligand d'une séquence polynucléotidique ou d'acides aminés par l'utilisation d'un substrat ayant une surface, ladite surface comprenant au moins  $10^3$  régions prédéfinies. lesdites régions prédéfinies contenant différentes séquences nucléotidiques ou d'acides aminés, lesdites régions prédéfinies occupant chacune une aire inférieure à environ  $2,5 \times 10^{-3}$  cm<sup>2</sup>, ledit procédé comprenant le fait de marquer ladite séquence soumise à un investigation et le fait d'identifier laquelle des différentes séquences se lie avec ladite séquence analysée.

2. Procédé selon la revendication 1, dans lequel ladite aire est inférieure à  $10\ 000\ \mu\text{m}^2$ , éventuellement inférieure à environ  $100\ \mu\text{m}^2$ .

3. Procédé selon la revendication 1 ou la revendication 2, dans lequel il y a  $10^4$ , ou davantage, séquences de monomères différentes et régions prédéfinies, éventuellement  $10^5$  ou davantage, sur ladite surface.

4. Procédé selon l'une quelconque des revendications 1 à 3, dans lequel ledit marquage est un marquage fluorescent, un marquage radioactif ou un anticorps marqué.

5. Procédé selon l'une quelconque des revendications 1 à 4, dans lequel les séquences de monomères dans les régions prédéfinies où une liaison est détectée, sont utilisées pour déterminer toute ou partie d'une séquence complémentaire à la séquence analysée.

6. Procédé selon l'une quelconque des revendications 1 à 5, dans lequel les séquences de monomères ont de 2 à 20 monomères de long.

7. Dispositif pour investigation par liaison récepteur/ligand d'une séquence polynucléotidique ou d'acides aminés comprenant un substrat ayant une surface, ladite surface comprenant au moins  $10^3$  régions prédéterminées, lesdites régions prédéterminées contenant différentes séquences polynucléotidiques ou d'acides aminés, lesdites régions prédéfinies occupant chacune une aire inférieure à environ  $2,5 \times 10^{-3}$  cm<sup>2</sup>,

8. Dispositif selon la revendication 7, qui est défini en Outre par la caractéristique spécifique de l'une quelconque ou plusieurs des revendications 2, 3, ou 6.

9. Dispositif selon la revendication 7 ou la revendication 8, dans lequel ledit substrat a une épaisseur inférieure à 0,5 mm, éventuellement inférieure à 0,05 mm.

10. Dispositif selon l'une quelconque des revendications 7 à 9, dans lequel lesdites régions prédéfinies ont une aire de surface inférieure à environ  $10^{-6}$  cm<sup>2</sup>.

11. Dispositif selon l'une quelconque des revendications 7 à 10, dans lequel chaque séquence de monomères d'acides aminés ou polynucléotidique est pure au moins à 50%, de préférence au moins à 90%, dans une région prédéfinie.

12. Utilisation d'un dispositif tel que défini selon l'une quelconque des revendications 7 à 11, pour l'investigation par



**EP 0 619 321 B1**

liaison récepteur/ligand d'une séquence polynucléotidique ou d'acides aminés.

5

10

15

20

25

30

35

40

45

50

55

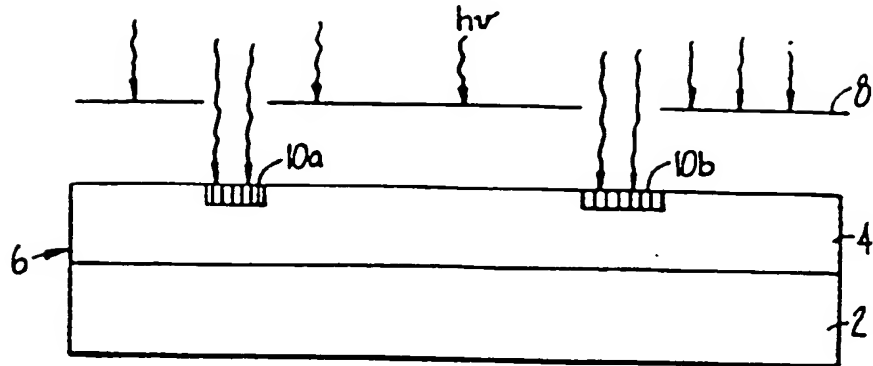


FIG. 1.

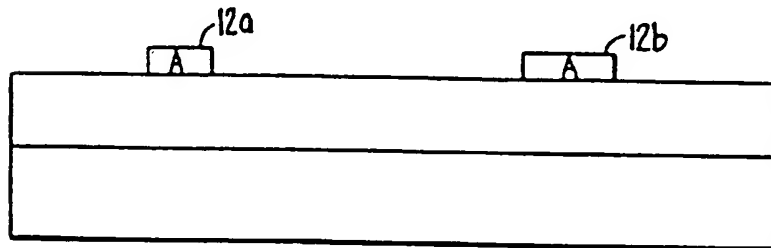


FIG. 2.

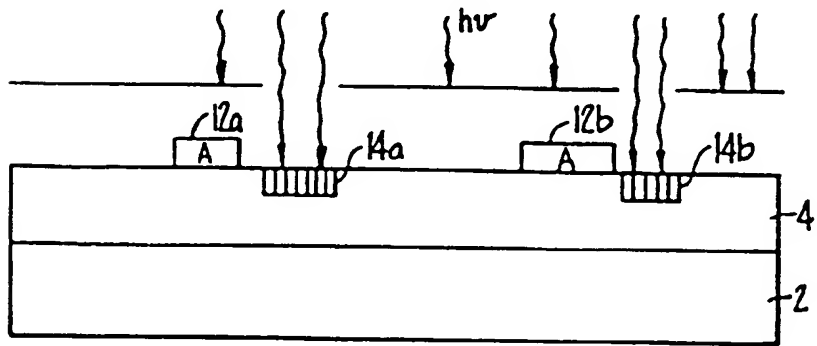


FIG. 3.

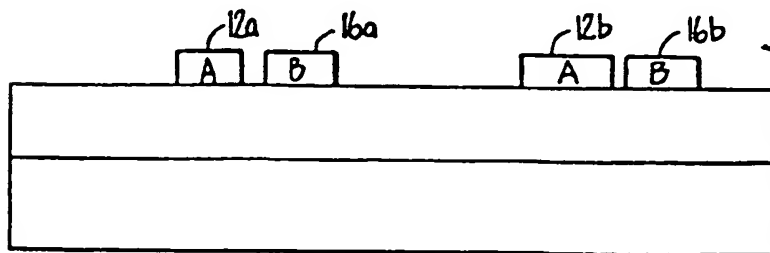


FIG. 4.

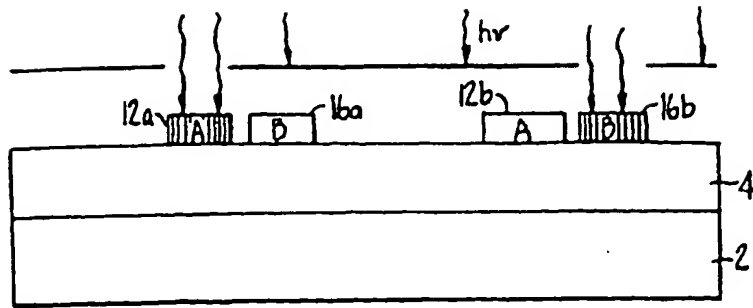


FIG. 5.

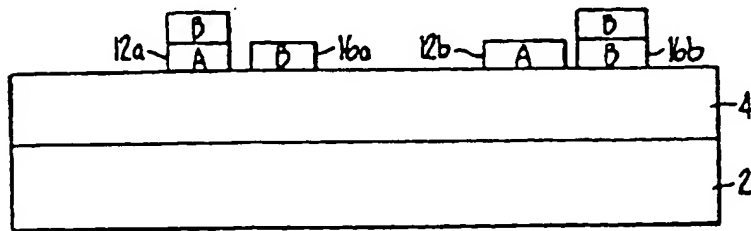


FIG. 6.

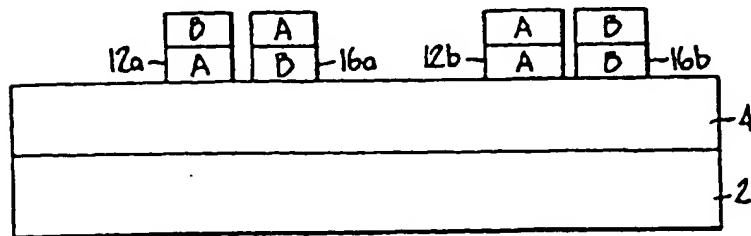


FIG. 7.

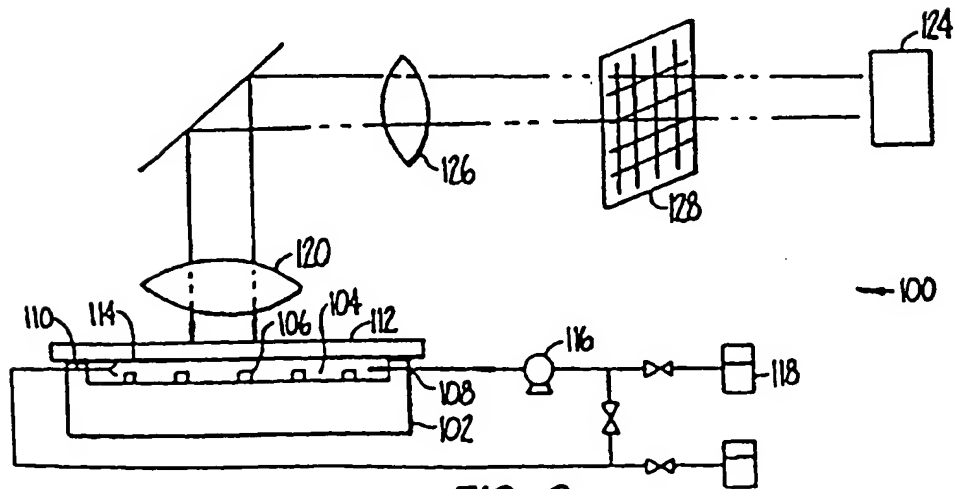


FIG. 8a.

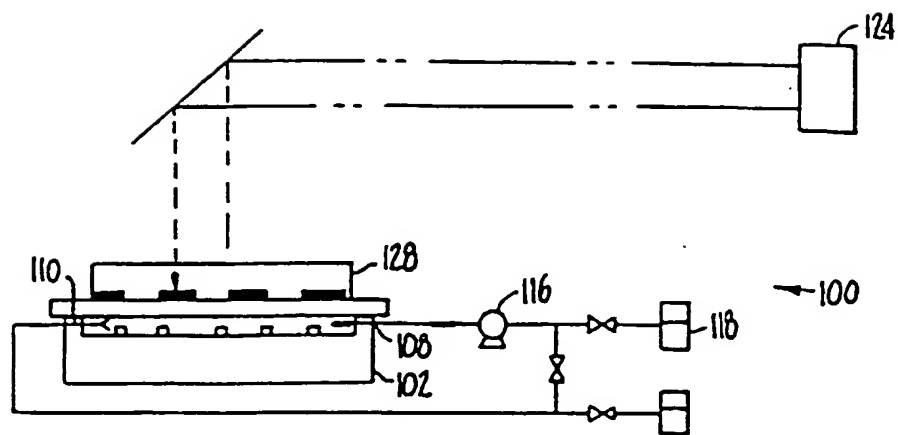


FIG. 8b

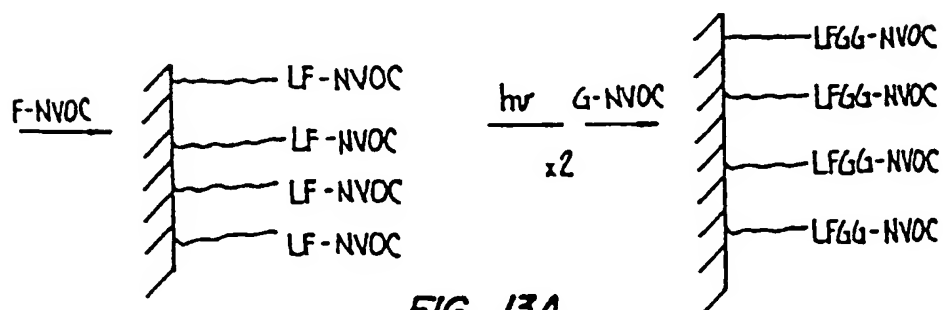
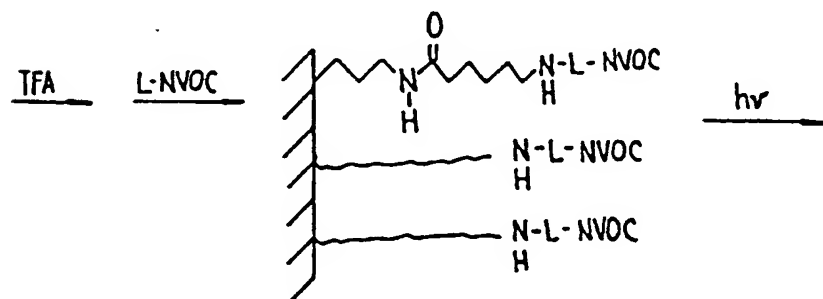
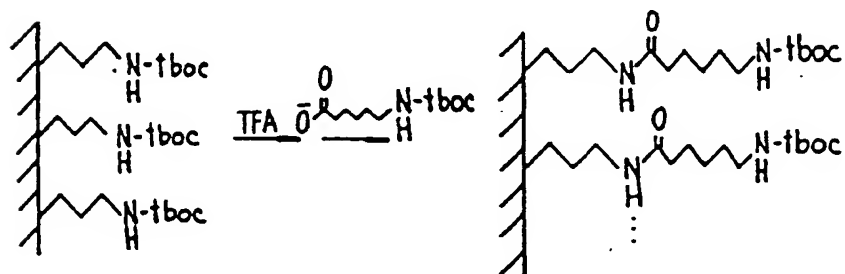


FIG. 13A.

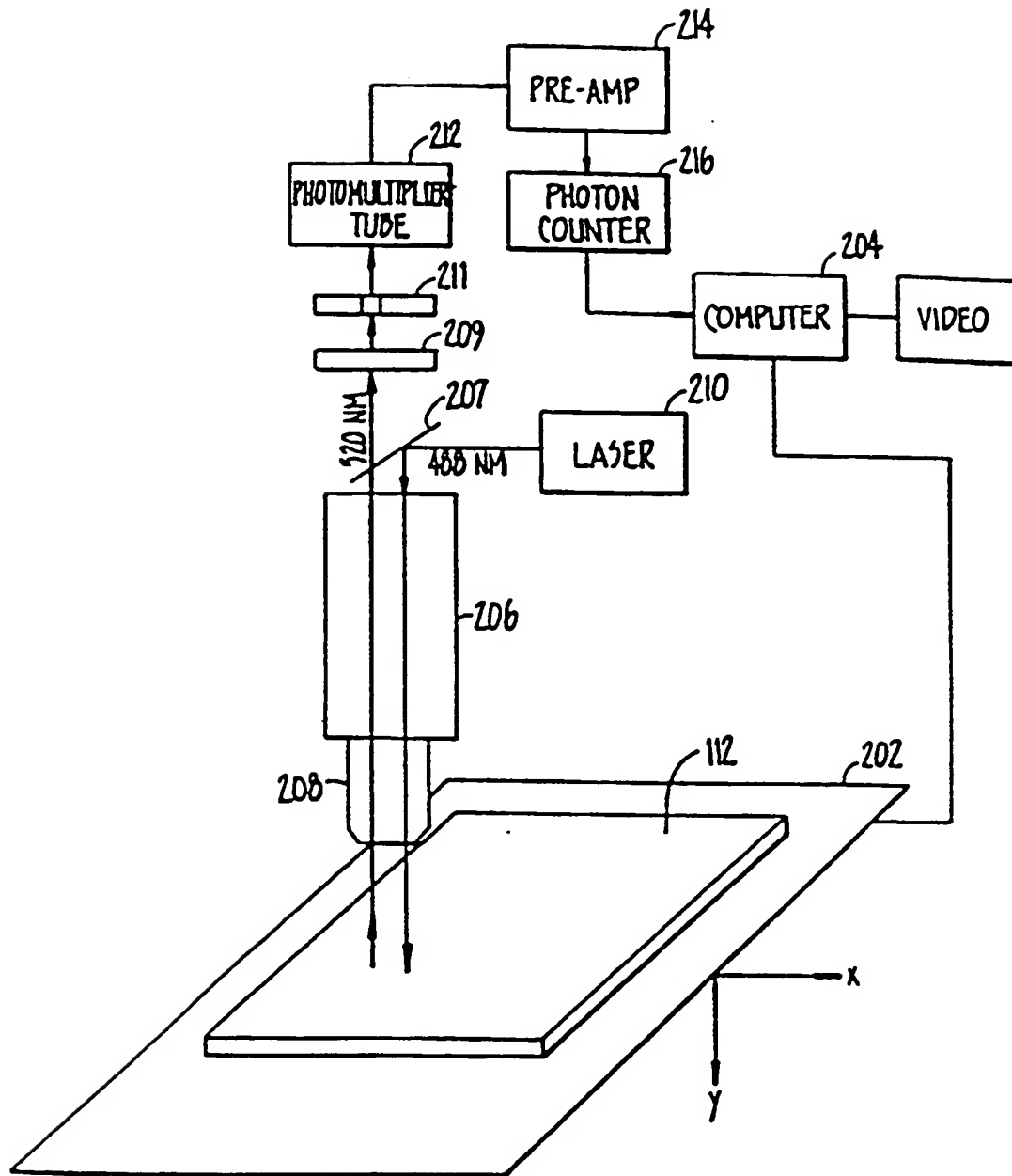


FIG. 9.

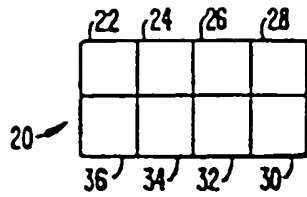


FIG. 10A.

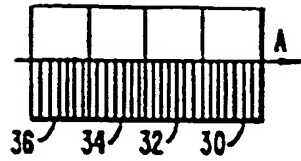


FIG. 10B.

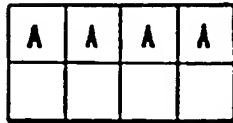


FIG. 10C.

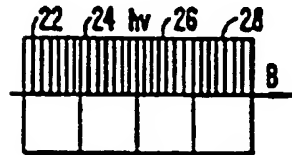


FIG. 10D.

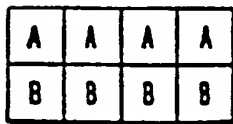


FIG. 10E.

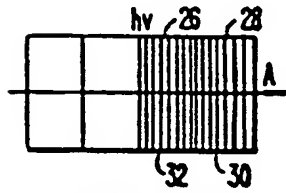


FIG. 10F.

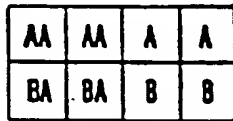


FIG. 10G.

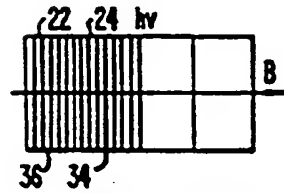


FIG. 10H.

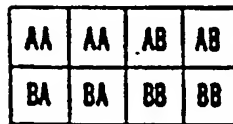


FIG. 10I.

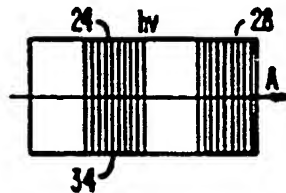


FIG. 10J.

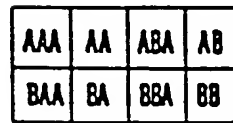


FIG. 10K.

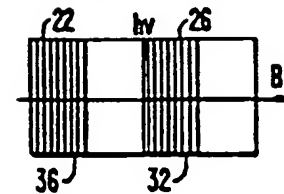


FIG. 10L.

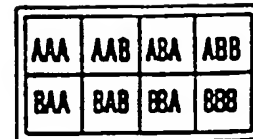


FIG. 10M.

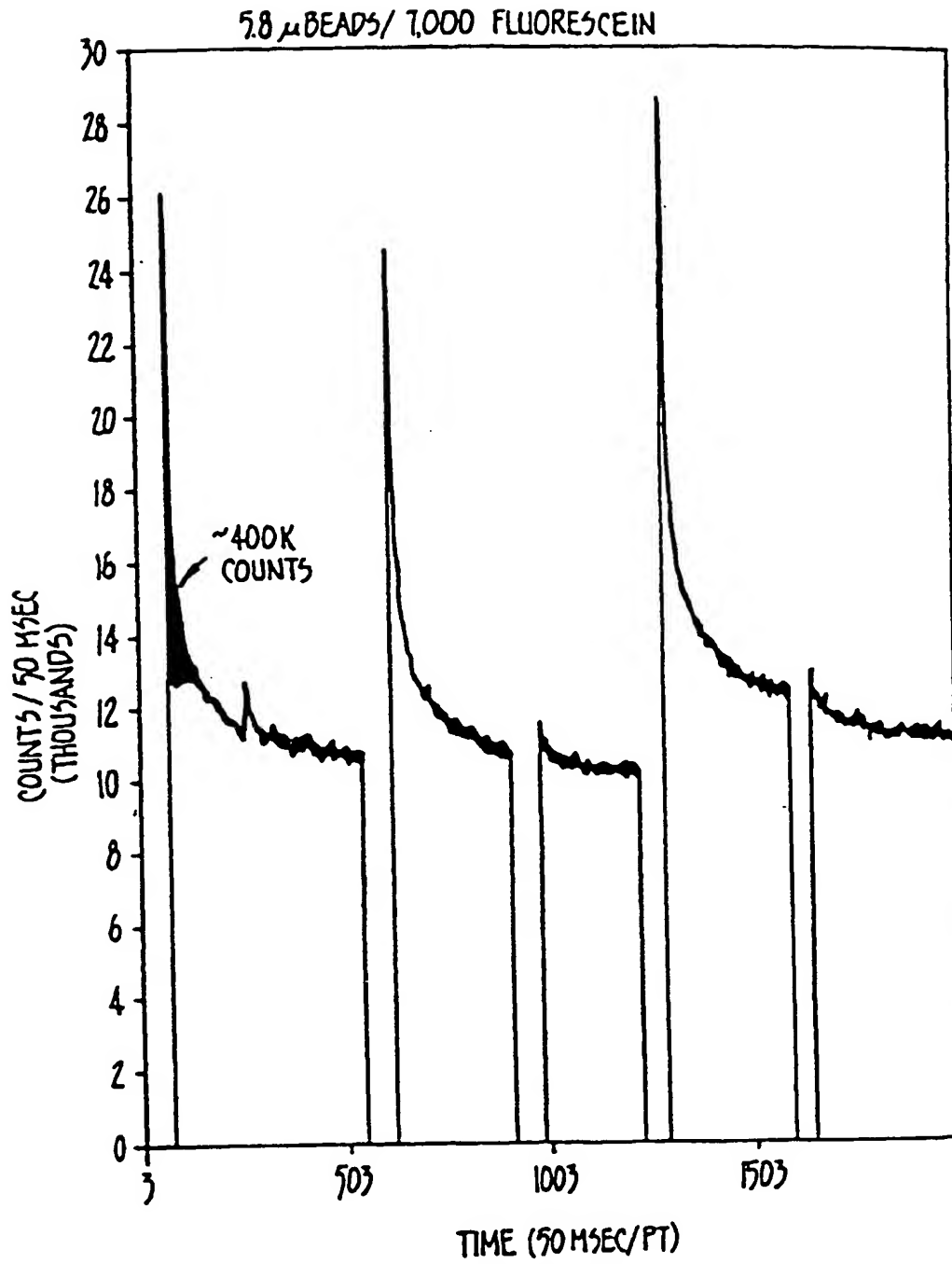


FIG. 11A.

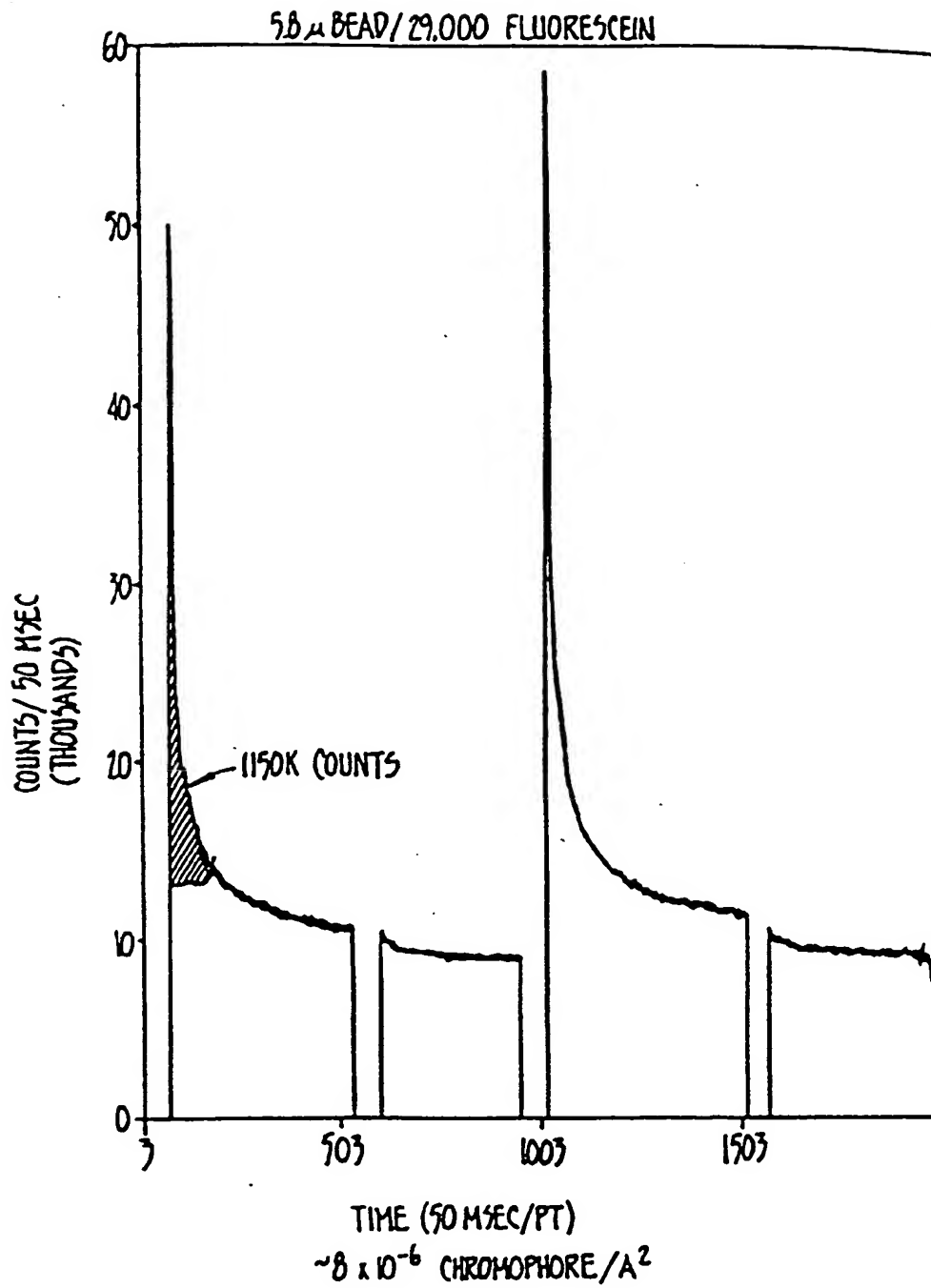
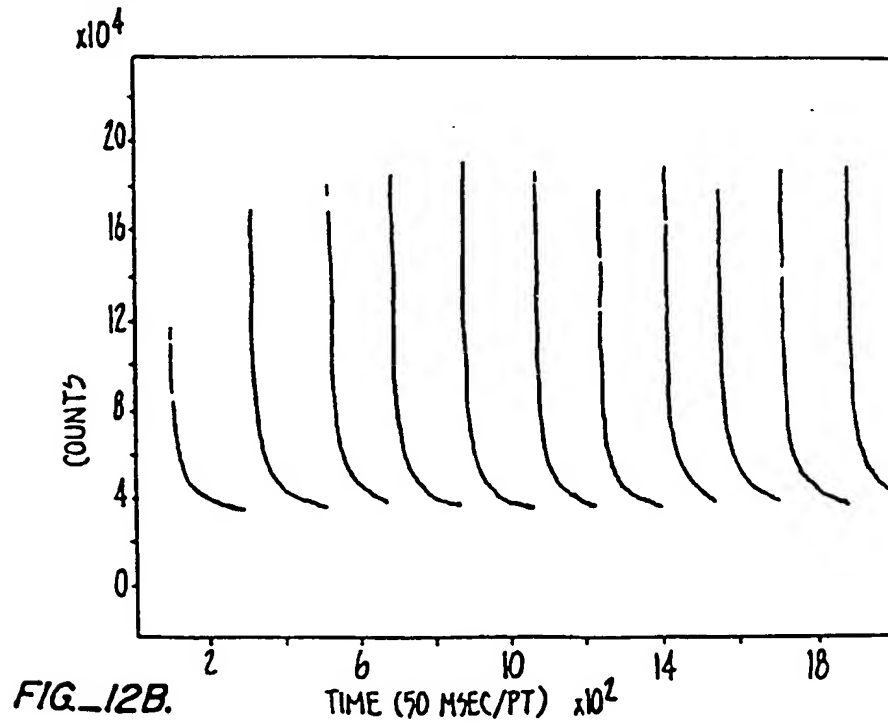
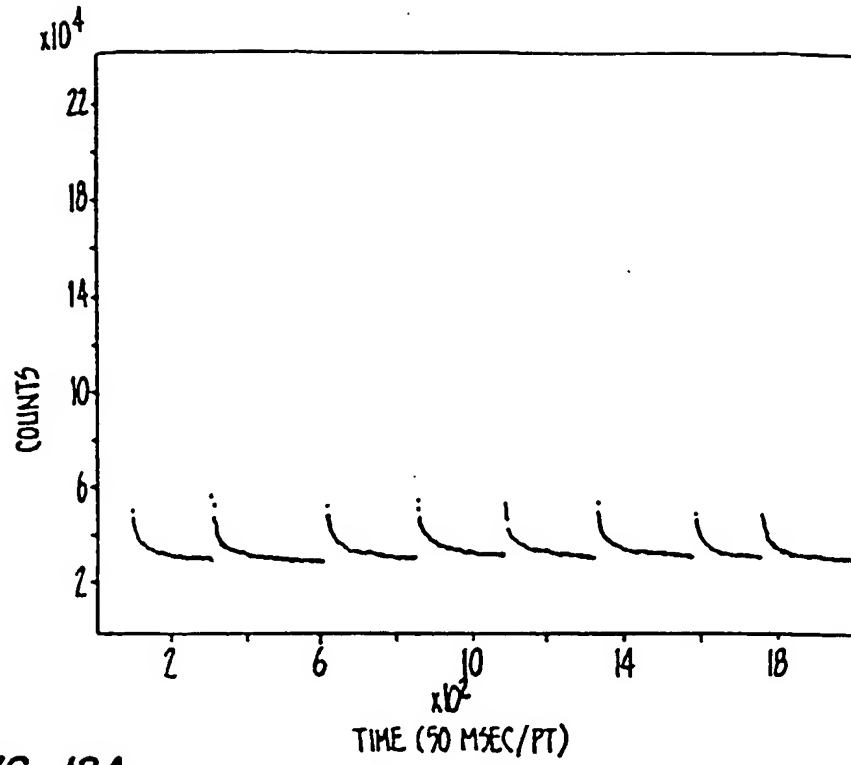


FIG. 11B.





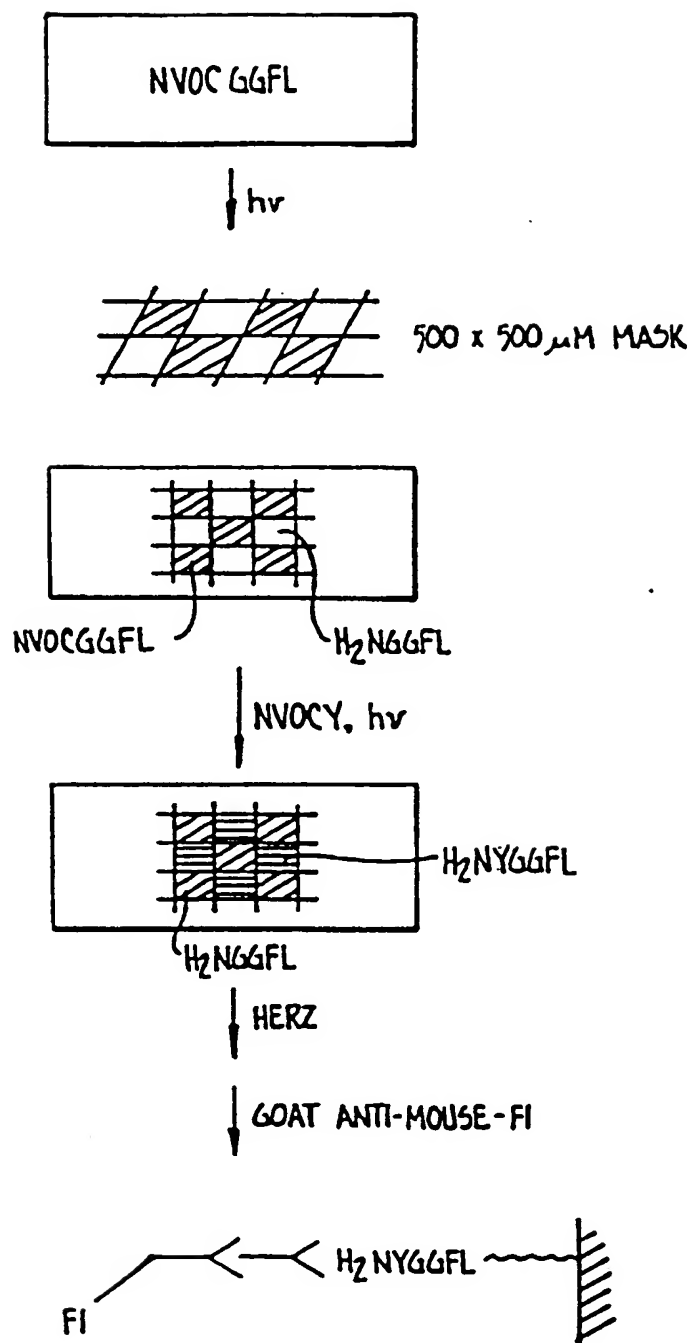


FIG. 13B

P	A	S	G	
<u>LP</u> GL	<u>LA</u> GL	<u>LS</u> GL	<u>LG</u> GL	L
<u>FP</u> GL	<u>FA</u> GL	<u>FS</u> GL	<u>FG</u> GL	F
<u>WP</u> GL	<u>WA</u> GL	<u>WS</u> GL	<u>WG</u> GL	W
<u>YP</u> GL	<u>YA</u> GL	<u>YS</u> GL	<u>YG</u> GL	Y

L SET

FIG. 14A.

P	a	s	G	
<u>Yp</u> GL	<u>Ya</u> GL	<u>Ys</u> GL	<u>YG</u> GL	Y
<u>fp</u> GL	<u>fa</u> GL	<u>fs</u> GL	<u>fg</u> GL	f
<u>wp</u> GL	<u>wa</u> GL	<u>ws</u> GL	<u>wg</u> GL	w
<u>yp</u> GL	<u>ya</u> GL	<u>ys</u> GL	<u>yg</u> GL	y

D SET

FIG. 14B.

**THIS PAGE BLANK (USPTO)**